# Characterizing Brain White Matter with Diffusion-Weighted Magnetic Resonance

Von der Fakultät für Physik und Geowissenschaften
der Universität Leipzig
genehmigte

DISSERTATION

zur Erlangung des akademischen Grades
Doctor rerum naturalium
Dr. rer. nat.

vorgelegt von

A.B, Bibek Dhital
geboren am 7. Februar 1984 in Nepal

Gutachter

Prof. Dr. Robert Turner

Prof. Dr. Daniel Alexander

Leipzig, den 24.04.2015

# *Bibliographische Beschreibung*

Dissertation von Bibek Dhital, Universität Leipzig

**Characterizing Brain White Matter with Diffusion-Weighted Magnetic Resonance**

170 Seiten, 257 Literaturangaben, 47 Abbildungen, 3 Tabellen

It has been known for almost two decades that the water proton NMR signal of diffusing water molecules in brain white matter undergoes a non-monoexponential decay with increasing diffusion gradient factor b. With the help of numerical simulations and analytical expressions, much effort has been directed to describing the signal decay and to extracting relevant biophysical features of the system under investigation. However, the physical basis of such nonmonoexponential behavior is still not properly understood.

The primary difficulty in characterizing this phenomenon is the variation in behavior in the different directions of diffusion measurement. A combined framework that accounts for the diffusion process in all directions requires several parameters. Addition of many such parameters renders a model to be unwieldy and over-complicated, but over-simplifications can be shown to miss crucially relevant information in the data.

In this thesis, I have attempted to handle this problem with simple measurements that span a wide range of parameter space. Compared to often-performed measurements that probe diffusion over a time-scale of 50-100 ms with relatively low diffusion weighting, the measurements here have been done for very short diffusion times of 2 ms and also very long diffusion times up to 2 s. The temperature dependence of the diffusion coefficients has also been extensively probed. To avoid problems related to gross tissue heterogeneity, diffusion-weighted MR imaging in vivo was performed with ultra-high resolution. These simple measurements allowed sequential assessment of many possible arguments that could have led to such non-monoexponential decay curves. Finally, it was concluded that the water in the glial processes was the major contributor to the non-exponential decay, giving rise to a 'slow' component both along the axonal fibers and transverse to them.

# Contents

# Chapter 1

# Introduction

The study of water diffusion holds a great potential in understanding structure and chemistry of its surrounding environment. Nuclear magnetic resonance (NMR) offers a unique chance to study water as a means to understand biological systems. The abundance of water in biological tissue means high signal-to-noise ratio (SNR) and non-invasive nature of the NMR probe allows us to study complex tissues with minimum experimental interference. In a diffusion NMR experiment the water molecules are tagged and the final signal is recorded as function of their displacement. The displacement of the molecules would depend on the microscopic determinants of diffusion like temperature and pH but also on the structure of the tissue that the molecules reside in.

Water diffusion in ideal systems with no structure in the near vicinity can be completely described by a single diffusion coefficient. Real biological systems are crowded with many macromolecules residing in a rather limited space. The brain white matter (WM) is one such complex, crowded environment.

Diffusion NMR reveals only an ensemble average description of displacement for all water molecules residing in our region of interest. To understand the collective property, it is necessary to know how microscopic substructures relate to the whole of WM. However, although the structural units that make up the WM has been studied in detail, the presence of water has largely been ignored. Even until few decades ago, water was treated as a liquid that homogeneously fills up the void space between macromolecules with no specific role of its own. The understanding of water dynamics in WM is still not well established.

The primary goal of this thesis was to understand the link between water diffusion in brain WM and the signal obtained from a diffusion NMR measurement. The majority of the thesis deals with changes in experimental parameters and their effects in the observed signal. Some part of the experimental work is also focused on optimal experimental design and data analysis.

Below is a short outline of the subsequent chapters.

## Outline

**Chapter 2** presents the background on diffusion physics. It is an overview on the description of diffusion process and how it is influenced by various conditions.

**Chapter 3** contains description of the NMR technique. It starts with the requisite background on NMR physics, signal generation and the influence of the environment on different MR parameters. This chapter also discusses on how diffusion NMR measurement is performed and how changing experimental parameters and sample properties affect the NMR signal.

**Chapter 4** contains two parts. The first part is a detailed description of the composition of the WM. It describes various substructures and tries to elucidate how they could affect diffusive behavior of water in and around those structures. The second part of this chapter is a review of the previous efforts in explaining the characteristics of the diffusion NMR signal.

Chapters 5, 6 and 7 form the experimental part of this thesis. They include brief backgrounds on the specific measurements followed by the measurements and conclusions that could be drawn from the results.

**Chapter 5** contains experiments performed on freshly excised corpus callosum of a pig. In this experiment, I obtained samples from the local slaughterhouse within 15 minutes of sacrifice and measured in a NMR spectrometer with very strong (35 T/m) magnetic gradients. This allowed measurements to be done at a very short diffusion time of 2 ms. Measurements at such short times allow fewer troubles due to relaxation and exchange between different compartments. Measurements were also performed at different temperatures. Additionally, CPMG and FIDs were also recorded.

The inverse Laplace transform of the diffusion measurement could distinguish two major diffusion components. Temperature dependence on the volume fraction and diffusion coefficients of the two components forms the major part of this chapter. The possible sources for the measured activation energy of diffusion have been discussed.

At the freezing transition, this measurement also showed a good correspondence between the unfrozen water fraction and the myelin water. Although not related to diffusion this is perhaps the first direct evidence of myelin water with its spin-spin relaxation behavior.

**Chapter 6** lays the groundwork for *in vivo* measurements. The MR scanner available for *in vivo* human measurements have a far weaker gradient system which allows diffusion measurements to be performed only at long observation times. Because of this limitation, in these systems it is perhaps better to understand diffusion at long time.

The first part of this chapter deals with issues related to measuring diffusion at high fields. The most pressing issues are shorter relaxation times and high specific absorption rate (SAR). Because of these issues and the need to measure at long diffusion times, stimulated echoes are probably the best way to acquire data in human scanners. I have further described the modifications made to stimulated echo sequence based on the current needs.

The second part of the chapter contains four experiments.

The first experiment is performed with diffusion tensor imaging (DTI) for a range of diffusion times between 50 and 2000 ms. The DTI measurement questions diffusion in WM as being fractal in nature. The time dependence of the tensor eigenvectors and the approach to Gaussian diffusion have been discussed.

The second experiment is a comparison between different fitting methods to the diffusion data. We find that for single direction measurement biexponential decay is probably the best way to fit the data. However, for multiple directions biexponential decay fails. The implication of these measurements have been discussed. A section of the discussion also deals with why fiber dispersion could not be a major source of non-monoexponential decay in WM.

The third experiment explores ways to increase the reliability and efficiency of diffusion measurements both in acquiring data and its analysis. As a side result, the analysis also showed that the physiological noise in MRI is mostly contributed by the cerebrospinal fluid (CSF). The CSF signal can be suppressed using diffusion weighting. This method can also be applicable in other MR modalities like BOLD which often is limited by the physiological noise.

The last experiment is performed at long times with strong diffusion weighting. The effect of long diffusion time measurement and exchange between diffusion compartments is evaluated for the diffusion decay curves. In this chapter I establish that the glial processes are the most likely cause of non-monoexponential decay in all directions, with axons also adding to the 'slow' component mix in the parallel direction.

**Chapter 7** is a discussion on the physical basis of diffusion in brain WM and a revisit on different compartments that water resides in. The problem of exponential decays has been looked again. Some ways to further check the findings in this thesis have been discussed. Finally, the implication of the findings on future modeling and current model selection have been discussed.

# Chapter 2

# Physics of Diffusion

This chapter introduces diffusive process as a ubiquitous phenomenon and describes its behavior in response to different physical conditions. It starts with a brief introduction of Einsteinian description of diffusion as a sum over random trajectories followed by a mathematically equivalent description derived from Fick's laws. The diffusion coefficient and its relation to temperature, pressure, viscosity and local geometry has been briefly discussed. The diffusion-geometry relationship has been explored. The relationship between the apparent diffusion coefficient and the ensemble average propagator is discussed in a heterogeneous medium with multiple compartments.

## 2.1 Diffusion as Random Motion

Diffusion is matter transport due to random thermal motion. It differs from inertial motion in that for diffusive motion, the displacement of an individual particle at any time is unpredictable; independent of its own history and the displacement of other particles. In this sense, for classical mechanics diffusion is effectively error or noise.

Following Brown's [1] observation of randomly moving pollen grains on water surface, the first description of sub-microscopic particles with persistent random motion was provided by Einstein [2]. The microscopic motion can be seen as a summation of discrete jumps such that the jump in any direction is independent of the jumps in orthogonal directions. For a given direction, $\vec{x}$, each jump has equal probability of jumping left or right. After $n$ jumps, the total displacement, $X$, in x-axis is given by

$$X(n) = \sum_{i=1}^{n} \varrho_x \ \alpha_i \tag{2.1}$$

where $\varrho_x$ is the mean distance per jump and $\alpha_i$ is a random binary number that could be $\pm 1$ depending on left or right jump. Mathematically, this describes a binomial distribution with equal probabilities for success and failure. If the process is repeated $n$

times, the expected outcome (mean) is zero and the variance $n\varrho_x^2$. Since these displacements happen in very small time steps the macroscopic outcome is cumulative of large number of such small motions. As $n$ gets bigger, de Moivre's [3, 4] theorem states that the binomial distribution asymptotically approaches the normal distribution (Gaussian) and a Gaussian with zero mean can be described with just its variance, $\sigma^2$, alone. The variance increases linearly with time. To be able to think of the diffusivity as an intrinsic property, the diffusion coefficient, $D$, is taken to be proportional to the ratio between the variance and the observation time. In one dimension, $D = \sigma^2/2t = \langle x^2 \rangle/2t$, where $t$ is the observation time. It is related to the microscopic step by $D_x = \varrho_x^2/2\tau$. For isotropic diffusion in three dimensions $D(t) = \langle RR^T \rangle/6t$ where $R$ is the vector displacement at time $t$.

Einstein's picture of diffusive motion is intuitive and was the first scientific theory that explicitly required the existence of small sub-microscopic particles. Though the random motion of a single particle at any time is unpredictable, the random walk illustration statistically describes the cumulative motion in terms of the mean-squared displacement. In reality, the distribution of these elementary molecular steps could be very different from a binomial distribution, but after many iterations of such small steps, the net displacement can still be described by a Gaussian function. The convergence of all distributions with finite variance to a Gaussian is guaranteed by the central limit theorem.

## 2.2 Diffusion as Mass Transfer

Einstein's random walk provides an intuitive way to understand diffusion as a physical stochastic process, but half a century before Einstein, Fick has already started studying diffusion as a mass transfer process [5]. This arose from observations in fluids that if the initial concentration $c$ of a particle is not uniform, the difference in concentration between two locations decreases over time. Fick noticed that mass transfer due to concentration difference was analogous to heat transfer in the presence of a temperature gradient. Using equations derived by Fourier for heat transfer, diffusion as mass transfer process was first described by Fick.

According to Fick's first law, a flux $J$ of particles through a unit area is proportional to the spatial gradient of concentration $C$,

$$J = -D \nabla C \tag{2.2}$$

with the diffusion coefficient $D$ as the constant of proportionality. Fick's second law relates the rate of change in concentration to the spatial variation in the flux. Along

with the the Eq. 2.2, it gives us

$$\frac{\partial C}{\partial t} = -\nabla(J) = \nabla.(D\nabla C) \tag{2.3}$$

For one dimension and a constant D, the solution to Eq. 2.3 can be

$$C = \frac{A}{\sqrt{t}} e^{-x^2/4Dt} \tag{2.4}$$

where $A$ is an arbitrary constant. Here, $C$ is symmetric at $x = 0$ for all $t > 0$. At $t = 0$, it is zero at everywhere except at $x = 0$. To normalize the total quantity of the diffusing particles, we use Gaussian integral property that gives $A = 1/\sqrt{4\pi D}$. $C$ is equivalent to a Gaussian function with zero mean and variance $2Dt$.

Fick's laws picture diffusion only in the presence of a concentration gradient. In his description, difference in concentration causes a kind of 'diffusive pressure' that is the cause of the resulting diffusive motion. But, the microscopic Brownian motion as described by Einstein regards diffusion as ubiquitous and incessant motion that causes the leveling of the concentration gradient. However, both still describe the final density function to be a Gaussian. Notwithstanding causality implications, we can mathematically reconcile Fickian diffusion with Einstein's random walk description by considering a diffusion of labeled particles in a solution of unlabeled yet identical particles.

## 2.3 Relation to other physical parameters

As a stochastic process, diffusivity is directly affected by any external condition that impacts microscopic motion. An increase in temperature, for instance, increases the average kinetic energy and increases the diffusivity of the system. The Stokes-Einstein relation describes the relationship between temperature, diffusion and viscosity.

$$\frac{D_1}{D_2} = \frac{\Theta_1}{\Theta_2}\frac{\mu_2}{\mu_1} \tag{2.5}$$

where, $\Theta_1$ and $\Theta_2$ are the two temperatures in absolute scale (K), and $\mu$ and $D$ are the viscosity (dynamic) and diffusivity of the medium at those temperatures. Similarly diffusion is also related to pressure.

$$\frac{D_1}{D_2} = \frac{\rho_2}{\rho_1} \tag{2.6}$$

where $D_{1,2}$, and $\rho_{1,2}$ are diffusivity and the mass density at two different pressures.

The dependence of diffusion with temperature is usually expressed with the Arrhenius equation.

$$D = D_0 e^{-E_A/R\Theta} \tag{2.7}$$

where $D_0$ is the constant of proportionality, $R$ is the universal gas constant and $\Theta$ is the temperature. The Arrhenius relationship for solid diffusion is related to breaking of bonds and more of less empirically observed for liquids.

Geometrical boundaries offer physical restriction to the free path of the molecules and affect their diffusivity.

Any interaction that affects the microscopic motion affects the resultant diffusivity. This allows diffusion measurement to be a sensitive probe to detect tiny differences in the environment but also makes it non-specific.

## 2.4 Anisotropic Diffusion

In simplified situations it is fair to consider diffusion as isotropic, but, in many circumstances it need not be. For example if tiny iron particles are allowed to 'float' in a room, they should show bigger spread in north-south direction over east-west direction: although they do not prefer north over south, it is easier for them to move along the earth's magnetic field line rather than across it. Liquid crystals, due to their molecular structure, are widely known to show anisotropic diffusion [6]. In such a system, diffusion is anisotropic, Eq. 2.2 is still valid, but with D expressed as a tensor,

$$D = \langle RR^T \rangle / 6t = \begin{pmatrix} D_{11} & D_{12} & D_{13} \\ D_{21} & D_{22} & D_{23} \\ D_{31} & D_{32} & D_{33} \end{pmatrix} \tag{2.8}$$

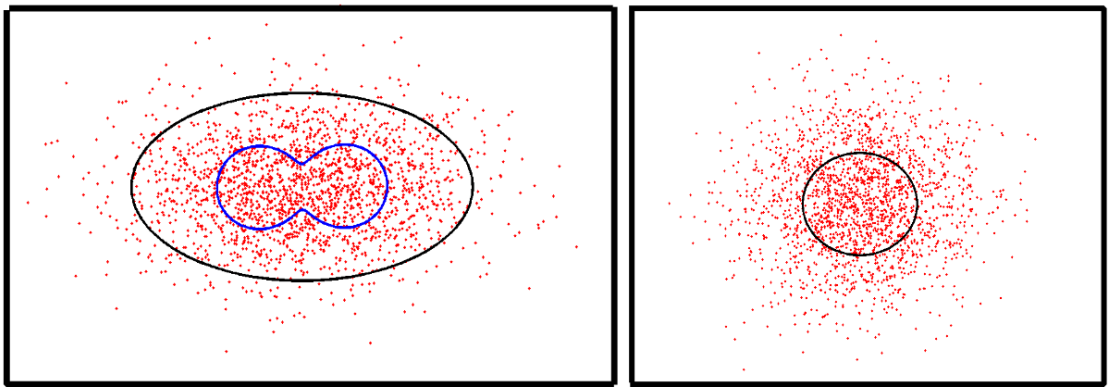


FIGURE 2.1: Simulation of anisotropic (left) and isotropic (right) diffusion in 2D with 2000 particles and 10000 steps. The red dots are particle displacement. The blue solid line is the diffusion coefficient measured along each direction (scaled) and the black line in the scaled isoprobability curve (analytically drawn). For the isotropic case both the diffusion coefficient and the isoprobability curves are circles(overlapped on each other). For anisotropic case, the diffusion coefficients take the dumbbell shape and the isoprobability curve is an ellipse.

The diffusion tensor is symmetric and positive definite [7]. Since diffusion is a random process, the anisotropy cannot be due to any force field that produces a net drift, but

only by reaction forces - those that arise by virtue of Newton's third law of motion. Many such forces exist like drag force, force due to charge moving in magnetic field and forces due to Newton's third law when a particle hits a barrier. Macroscopically force fields are continuous and they can be easily incorporated to explain anisotropic diffusion [1].

Diffusion anisotropy can also be observed due to geometrical barriers. If the distance between the barriers is large, diffusion could be spatially dependent - Gaussian far from the walls but not in the vicinity of the walls. In some cases, the departure from Gaussianity could be transient: as molecules diffuse and move through different locations, central limit theorem ensures that after a sufficiently long time $t_\infty$ the diffusion will again be a Gaussian, albeit with a reduced diffusion coefficient that accounts for the effect of the barriers.

In many cases the time required to approach asymptotic Gaussianity could far exceed the observation time possible. When diffusion is not a Gaussian, $D$ as a single parameter is no longer sufficient to describe the diffusion phenomenon and requires us to resort back to the probability density function (PDF) of the particle displacements. Different fields in physics have referred to this PDF as the diffusion propagator, Greene's function and van Hove self-correlation function.

## 2.5 Diffusion Propagator

Eq. 2.4 described the diffusion propagator for free diffusion in one dimension. The diffusion propagator, $P(\vec{\mathbf{r}} \,|\, \vec{\mathbf{r_0}}, t)$, is the probability that after time $t$ a particle starting at $\vec{r_0}$ is within $\vec{\mathbf{r}} \pm d\vec{\mathbf{r}}$. For free diffusion the initial and final positions were degenerate and only the net displacement $\vec{r} - \vec{r_0}$ was relevant. However, in the presence of boundaries, the initial and final positions also have to be considered. For a finite particle, the initial probability of finding it at a given position is the Dirac delta function $P(\vec{\mathbf{r}} \,|\, \vec{\mathbf{r_0}}, 0) = \delta(\vec{\mathbf{r}} - \vec{\mathbf{r_0}})$. Considering the initial distribution of all particles as $C(\vec{\mathbf{r}}, 0)$, at time $t$ the final distribution $C(\vec{\mathbf{r}}, t)$ can be written as

$$C(\vec{\mathbf{r}}, t) = \oint C(\vec{\mathbf{r}}, 0)\, P(\vec{\mathbf{r}} \,|\, \vec{\mathbf{r_0}}, t)\; dr \tag{2.9}$$

The diffusion propagator also follows Fick's law giving

$$\frac{\partial}{\partial t} P(\vec{\mathbf{r}} \,|\, \vec{\mathbf{r}}_0, t) = \nabla.(D\, \nabla P(\vec{\mathbf{r}} \,|\, \vec{\mathbf{r}}_0, t)) \tag{2.10}$$

Many labeling techniques such as dyes, radioactive tracers and single molecule fluorescence spectroscopy probe diffusion by observing net particle displacements at different

[1] In particle physics they are also treated as discontinuous

times. In principle, the experimental setup allows these techniques to reconstruct the propagator for different times and positions. Nuclear magnetic resonance (NMR) however, probes diffusion properties of all the labeled molecules in a single step. For NMR based diffusion measurements the information obtained is on the spatial average of the propagator at all starting positions $p(\vec{\mathbf{r}})$ ,

$$EAP = \int p(\vec{\mathbf{r}})\, P(\vec{\mathbf{r}}(t) - \vec{\mathbf{r}}_0\,|\,, t) \tag{2.11}$$

where EAP is the ensemble averaged propagator.

The propagator and the EAP contain complete information of what could be obtained from diffusion measurements, however inferring properties of the medium from this information might not be trivial. The following section considers some simple cases on how the propagator behaves.

## 2.6 Non Gaussian Diffusion

### 2.6.1 Restricted Diffusion

Diffusion in a space enclosed within perfectly reflecting walls is the ideal case for restricted diffusion. The PDF can be obtained by solving the differential equation describing Fick's law with boundary condition. Here we will derive the PDF in two restricted systems, diffusion in infinite parallel plates for one dimension and diffusion perpendicular to the axis of an infinitely long cylinder for two dimensional case.

#### 2.6.1.1 Diffusion between parallel plates

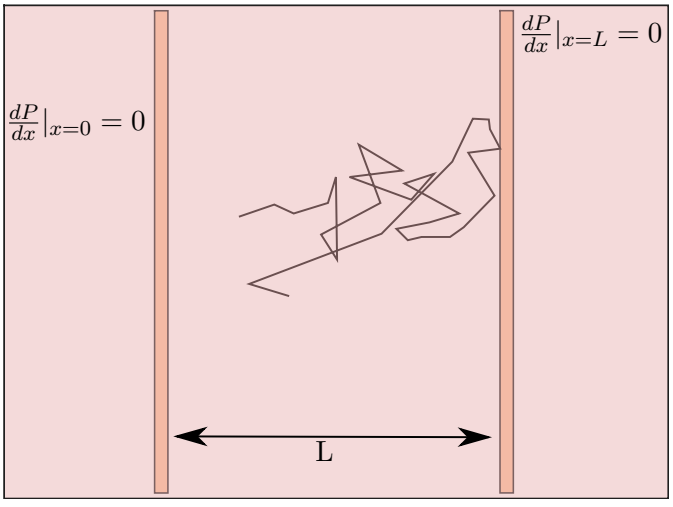



FIGURE 2.2: Diffusion inside parallel plates. The particles that hit the wall are reflected.

We will consider one-dimensional diffusion, bounded between two infinite, impermeable planes at $x = 0$ and $x = L$. To account for reflection at the boundaries we can equivalently think that the number of particle going 'out' through these boundaries is always equal to the number of particles simultaneously coming in so that the net flux is zero. The net concentration at the boundaries remains constant. Applying Fick's law to the diffusion propagator.

$$D\frac{\partial^2}{\partial x^2}P(x|x_0,t) = \frac{\partial P_s}{\partial t} \tag{2.12}$$

the initial and the boundary conditions are

$$P(x|x_0,0) = \delta_{x,x_0} \tag{2.13}$$

$$\left.\frac{\partial P(x|x_0,t)}{\partial t}\right|_{x=(0\vee l)} = 0 \tag{2.14}$$

Considering a separation of variables, $P = X(x)T(t)$, Eq. 2.12 can be transformed to

$$\frac{1}{X}\frac{d^2X}{dx^2} = \frac{1}{DT}\frac{dT}{dt} = -\lambda^2 \tag{2.15}$$

where $\lambda$ is a constant. The solution for such an equation can be written as

$$T(t) = A\ exp(-\lambda Dt) \qquad \text{and}$$

$$X(x) = B\ sin(\lambda x) + C\ cos(\lambda x) \tag{2.16}$$

Considering the boundary conditions (Eq. 2.14), at $x = 0$ we obtain

$$0 = \left.\frac{\partial P}{dx}\right|_{x=0} = T(t)\left.\frac{dX}{dx}\right|_{x=0} = T(t)B \Rightarrow B = 0$$

and for $x = L$,

$$0 = T(t)C\lambda sin(\lambda x) \Rightarrow \lambda = \frac{n\pi}{L}$$

where $n = 0, 1, 2..$ is an integer. Applying the initial condition (Eq. 2.13), we get

$$\delta_{x,x_0} = P(x|x_0,0) = X(x) = \sum_{n=0}^{\infty}\left[C_n cos\left(\frac{n\pi x}{L}\right)\right]$$

which is a Fourier cosine series for $\delta_{xx_0}$ on $0 \le x \le L$. This can be used to deduce the $C_n$s, giving

$$C_n = \begin{cases} \frac{1}{L}\int_0^L \delta_{x,x_0}dx = \frac{1}{L}\ , & n = 0 \\ \frac{2}{L}\int_0^L \delta_{x,x_0} cos\left(\frac{n\pi x}{L}\right)dx = \frac{2}{L}cos\left(\frac{n\pi x_0}{L}\right)\ , & n \neq 0 \end{cases} \tag{2.17}$$

The diffusion propagator $P(x|x_0,t)$ can be written as

$$P(x|x_0,t) \;=\; \frac{1}{L} + \frac{2}{L}\sum_{n=1}^{\infty}\left[cos\left(\frac{n\pi x}{L}\right) cos\left(\frac{n\pi x_0}{L}\right) exp\left(-\frac{n^2\pi^2}{L^2}Dt\right)\right] \tag{2.18}$$

#### 2.6.1.2 Diffusion perpendicular to infinite cylinder

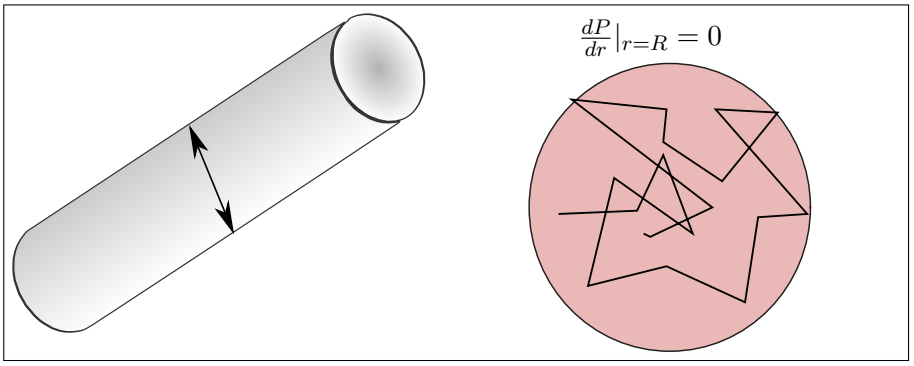



FIGURE 2.3: Diffusion inside restricted infinite cylinder. Diffusion is free along the axial direction and restricted in the radial direction.

Diffusion perpendicular to a cylinder is befitting to be considered in cylindrical coordinates where $\nabla^2 P$ takes the form,

$$\nabla^2 P = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial P}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2 P}{\partial \varphi^2} + \frac{\partial^2 P}{\partial z^2}. \tag{2.19}$$

where $[r, \varphi, z]$ are the basis set representing the radius, angle and height in cylindrical coordinates. For diffusion perpendicular to the infinite cylinder, z-axis is not relevant and Fick's second law can be written as

$$\begin{aligned}\frac{1}{D}\frac{\partial P}{\partial t} = \nabla^2 P &= \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial P}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2 P}{\partial \varphi^2} \\ &= \frac{\partial^2 P}{\partial r^2} + \frac{1}{r}\frac{\partial P}{\partial r} + \frac{1}{r^2}\frac{\partial^2 P}{\partial \theta^2}\end{aligned} \tag{2.20}$$

Similar to the previous case, using separation of variables we can consider $P(\vec{r}|\vec{r_0},t) = R(r)\Theta(\theta)T(t)$, and rearrange Eq.2.20 to solve for $T(t)$ and $\Theta(\theta)$, such that

$$\frac{1}{DT}\frac{dT}{dt} = \left[\frac{1}{R}\frac{d^2R}{dr^2} + \frac{1}{Rr}\frac{dR}{dr} + \frac{1}{\Theta r^2}\frac{d^2\Theta}{d\theta^2}\right] = -\alpha^2 \tag{2.21}$$

$$\frac{1}{\Theta}\frac{d^2\Theta}{d\theta^2} = -\lambda^2 \tag{2.22}$$

and

$$\left[\frac{d^2R}{dr^2} + \frac{1}{r}\frac{dR}{dr} + R\left(\frac{1}{r^2\Theta}\frac{d^2\Theta}{\theta^2} - \frac{1}{DT}\frac{dT}{dt}\right)\right] = 0 \tag{2.23}$$

where $\alpha$ and $\lambda$ are constants. The general solutions for $T$ and $\Theta$ is similar to that for $T$ and $X$ in the parallel plane case (Eq.2.16).

$$T(t) = A\ exp(-\alpha D t) \qquad \text{and}$$

$$\Theta(\theta) = B_1\ sin(\lambda\theta) + B_2\ cos(\lambda\theta) \tag{2.24}$$

Apart from trivial solutions, since $\theta$ is periodic at $2\pi$, $\sqrt{\lambda} = n$ are solutions to the equation above when $n$ is an integer.

And substituting Eq.. 2.21 and 2.22 to Eq. 2.23 gives

$$\left[\frac{d^2R}{dr^2} + \frac{1}{r}\frac{dR}{dr} + \left(\alpha - \frac{n^2}{r^2}\right)R\right] = 0 \tag{2.25}$$

which is akin to the Bessel equation of order $n$, the solution to which is should be [8]

$$R(r) = C_1 J_n(\alpha\, r) + C_2 Y_n(\alpha\, r) \tag{2.26}$$

Where $J_n$ and $Y_n$ are the order $n$ Bessel functions of the first and the second kind respectively. Since $Y_n$ is not finite at the origin, it is not a relevant solution here and can be ignored.

For the boundary condition, we have

$$0 = \left.\frac{\partial P}{\partial r}\right|_{r=R} = T\Theta\left.\frac{dR}{dr}\right|_{r=R} = C\alpha\, T\Theta J_n^{'}(\alpha R) \tag{2.27}$$

Except for trivial solutions, to satisfy the boundary condition any $\alpha$ must be a solution of $J_n^{'}(\alpha R) = 0$.

Combining solutions for all three independent variables, the diffusion propagator $P$ can be expressed as

$$P = A_0 \sum_k J_0(\alpha_k\, r) + \sum_k \sum_{n=1}^{\infty} exp(-\alpha_k^2 D t) J_n(\alpha_k\, r)(A_{nk} cos(n\theta) + B_{nk} sin(n\theta)) \tag{2.28}$$

or,

$$P(\vec{\mathbf{r}}\,|\,\vec{\mathbf{r}}_0, t) = \sum_{n=-\infty}^{\infty} \sum_k exp(-\alpha_{nk}^2 D t) J_n(\alpha_{nk} r)\,[A_{nk} cos(n\theta) + B_{nk} sin(n\theta)] \tag{2.29}$$

Using the Lommel's integrals (Eq. 2.47 in Appendix) and the recurrence relation (Eq. 2.45 in Appendix) $P(\vec{\mathbf{r}}\,|\,\vec{\mathbf{r}}_0, t)$ can be shown to be

$$P(\vec{\mathbf{r}} \mid \vec{\mathbf{r}}_0, t) = \sum_{n=-\infty}^{\infty} \sum_{k} \frac{exp(-D\alpha_k^2 t)}{\pi R^2 \left(1 - \frac{n^2}{\alpha R^2}\right)} \frac{J_n(\alpha_k r) J_n(\alpha_k r_0)}{\left[J_n(\alpha_k R)\right]^2} \times \left[cos(n\theta_0)cos(n\theta) + sin(n\theta_0)sin(n\theta)\right] \tag{2.30}$$

Now, Eq.2.30 along with the orthogonality conditions for sine and cosine results in the the solution for diffusion in the direction perpendicular to an infinite cylinder can be expressed as

$$P(\vec{\mathbf{r}} \mid \vec{\mathbf{r}}_0, t) = \frac{1}{\pi R^2} + \sum_{k} \sum_{n=1}^{\infty} \frac{2}{\pi R^2 \left(1 - \frac{n^2}{\alpha_k^2 R^2}\right) \left[J_n(\alpha_k R)\right]^2} exp(-D\alpha_k^2 t) \times J_n(\alpha_k r) J_n(\alpha_k r_0) \left[cos(n\theta_0)cos(n\theta) + sin(n\theta_0)sin(n\theta)\right] \tag{2.31}$$

The two examples above show that in a restricted systems the conditional probability $P(\vec{\mathbf{r}} \mid \vec{\mathbf{r}}_0, t)$ that a particle initially at $\vec{\mathbf{r}}_0$ is at $\vec{\mathbf{r}}$ after time $t$ is a function of both starting and end positions. We also see the symmetry such that probability of finding a particle that moved from location $\vec{\mathbf{r}}_0$ to $\vec{\mathbf{r}}$ is same as probability of finding a particle that started at $\vec{\mathbf{r}}$ and ended at $\vec{\mathbf{r}}_0$. We see that the function behaves as an exponential decay in time multiplied with the harmonics of the modeling expansion of the restricting space. At sufficiently long $t$ the exponential decay factor makes contributions from the harmonics to go to zero so that the only remaining term is the first term related to the density of the particles. At such long times, the probability of finding a particle at any position is independent of its initial position.

### 2.6.2 Heterogeneous Medium

So far we have seen that two cases: (i) In a homogeneous medium diffusion is Gaussian -the propagator at any time $t$ can be completely described by diffusion coefficient $D$. (ii) In a restricted medium the shape of the propagator is an exponential decay in time multiplied by the coefficients of the harmonics of the restricting medium geometry. At very short times it resembles a Gaussian, and at very long times it takes the shape of the self-convolution of the restricting space.

Free and restricted systems are only idealizations. Most naturally existing systems have different irregular shapes with different permeabilities. At short times, the asymptotic behavior of the diffusion coefficient in inhomogeneous media has been shown to be proportional to the surface-to-volume ratio of the medium [9–11]

$$D(t) \simeq D_0 \left[1 - \frac{4}{3d\sqrt{\pi}} \left(\frac{S}{V}\right) \sqrt{D_0 t}\right] \tag{2.32}$$

where $D_0$ is the bulk diffusion coefficient and $d$ is the effective spatial dimension. This formulation is only valid for very short times when the diffusion length ($\sqrt{2Dt}$) is very small, i.e. on average, the free diffusion trajectories sense only one boundary.

At long diffusion times the diffusion coefficient again approaches a constant value reduced by a proportionality constant. This constant of proportionality is referred as the tortuosity factor [11, 12]. In heterogeneous, disordered medium, diffusion time dependence is usually fitted with powers of $t^{-1/2}$ so that

$$\frac{D(t)}{D_0} \simeq \frac{1}{\alpha} + \frac{\beta_1}{t^{1/2}} + \frac{\beta_2}{t} + ..... \tag{2.33}$$

Here $\alpha$ is the tortuosity factor, and $\beta$s are the subsequent higher order terms that could be estimated. However, direct physical interpretations are hard to come by but they allow understanding of the type of disorder involved [11, 13, 14]. Recently, the approach to $D_\infty$ has been used to characterize the geometry of restricting boundaries [15].

Biological tissues are composed of structures with a broad range of size and shape: the tiniest of them are smaller than the length scales currently accessible in modern NMR systems. It is inevitable that the diffusion process is sensitive to obstructions in all scales and from all geometries. To emphasize the difference between our measured diffusion and the true diffusion process, diffusion coefficient is often referred as apparent diffusion coefficient (ADC) [16, 17]. In what follows we attempt to throw light on how the diffusion process is related to our observation in some simple yet analytically intractable system.

### 2.6.3 Flux through leaky walls

A simple case for non-ideal system is diffusion of water from inside of a leaky ball shown in Fig. 2.4. The water molecules that hit the boundaries are reflected most of the time but are allowed to escape with a probability $0 < p < 1$. A complete analytical solution is difficult even for such a simple system. Nevertheless, we can still understand what we observe at different times. When the barriers are quite permeable or 'weak', $0 << p < 1$, the molecules inside quickly go outside. In this situation the barriers only influence a small fraction of molecules that are close to the walls. At short times we can expect small deviation from the Gaussian behavior but as particles get out of the boundaries Gaussian behavior is quickly restored.

The problem gets more realistic when the barriers are 'strong', $0 < p << 1$. On average the particles hit the boundaries multiple times before escaping and $l_0 << \sqrt{D_0\tau_m}$, where $l_0$ is the length of the object, $D_0$ is the time independent bulk diffusion coefficient, and $\tau_m$ the mean residence time that a particle spends inside.

Let's first consider diffusion of water that originate from one such small region placed in a sea of bulk water. Here, the probability that a water molecule escapes the barriers

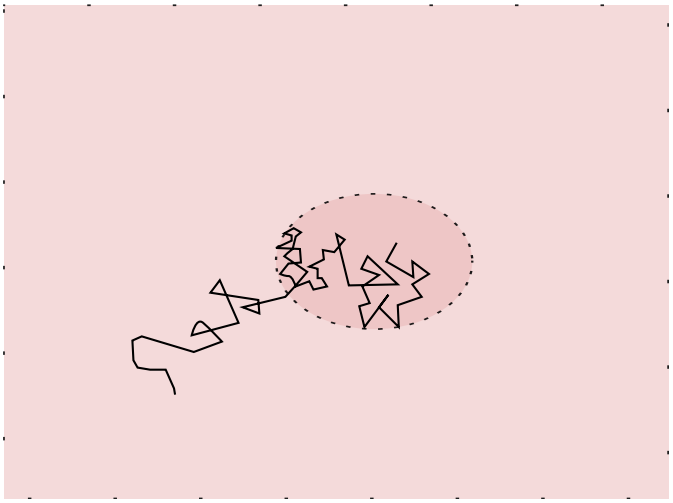


FIGURE 2.4: Diffusion through leaky walls.

at a given time is proportional to the number of particles remaining inside,

$$\frac{dP_i(t)}{dt} = -\frac{P_i(t)}{\tau_i} \Rightarrow \; P_i \;\; = \;\; exp^{-t/\tau_i} \tag{2.34}$$

$$\text{and} \quad P_o(t) \;\; = \;\; 1 - P_i(t) \tag{2.35}$$

where $P_i$ and $P_o$ are the probability that the particle is either inside or outside. Since the restricted space is very small, each molecule that escapes the walls has close to zero probability of re-entering the same region. At long times when $t > l_0^2/D$ , the mean squared displacement $\langle r^2 \rangle_i$ and $\langle r^2 \rangle_o$ can be approximated to the first order as

$$\frac{d\langle r^2 \rangle_i}{dt} = 0 \tag{2.36}$$

$$\frac{d\langle r^2 \rangle_o}{dt} = 2\mathrm{d}D_0 \tag{2.37}$$

$$\tag{2.38}$$

d being the dimensions. This can be solved as

$$\langle r^2 \rangle_i(t) = \langle r^2 \rangle_i(Res) \tag{2.39}$$

$$\langle r^2 \rangle_o(t) = \mathrm{d}D_0 t \tag{2.40}$$

where $\langle r^2 \rangle_i(Res)$ is the maximum attainable mean squared displacement within the barriers. It is the ensemble averaged propagator (EAP) assuming restricting boundaries. It should be emphasized that this is only valid for long times $t > l_0^2/\mathrm{d}D_0$. The ADC can be written as the weighted sum of the two diffusion coefficients $D_i$ and $D_o$ with relative

fraction $P_i$ and $P_o$

$$ADC(t) = \frac{\langle r^2 \rangle_i \times P_i + \langle r^2 \rangle_o \times P_o}{2\mathrm{d}t} = D_i(t) \times (1 - exp^{-t/\tau_i}) + D_0(t) \times exp^{-t/\tau_i} \quad (2.41)$$

The diffusion of molecules initially trapped between leaky walls is of interest for problems associated with drug delivery where the molecules are expected to reach the target at some desired rate [18]. But in many cases one is interested in the self-diffusion of the medium which is present both inside and outside the structures.

### 2.6.4 Diffusion in pores with leaky walls at steady state

When the system is in steady state the molecules keep entering and exiting the enclosed space with a zero net flux. Nevertheless, the residence time $\tau$ still helps to better describe the situation.

Every molecule that enters the region will on average spend time $\tau_m$ being inside. Reiterating that $l_0 << \sqrt{D_0 \tau_m}$, during the time spent inside, the molecule on average samples all of the pore space. Therefore, for the time between molecules enter and exit the pore we can approximate there ADC as the EAP$/2\mathrm{d}\tau_m$.

### 2.6.5 Diffusion in distributed pores with leaky walls

We can now further extend this gedanken to a more realistic situation where many small pores are floating in a sea of bulk water. For each region $r$, let's consider the mean residence time $\tau_m, r$ is far greater that the average time taken for the particles to sample the space, $\tau_m, r >> l_{0,r}^2/2dD_0$. The ADC for molecules entering and escaping each pore can still be ascribed with their own apparent diffusion coefficient $ADC_r$.

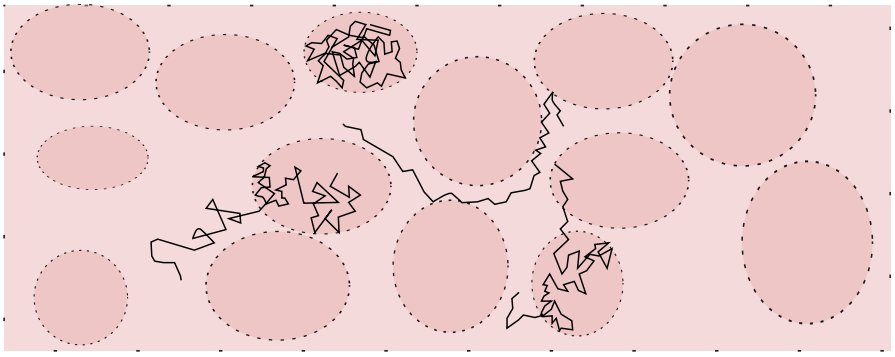


FIGURE 2.5: Diffusion in a multi-pore medium.

However, presence of many pores also affects diffusion in the bulk. The molecules that are outside these pores have to constantly face obstruction in their diffusion paths. Most of the time the barriers do not allow the molecules to pass through so molecules have to go 'around' the barriers following a tortuous path. At long times, the increased path

length reduces the mean squared displacement. Following Eq. 2.33 the hindered $ADC_h$ for the molecules in bulk is

$$ADC_h \simeq \frac{D_0}{\alpha} \tag{2.42}$$

In most circumstances $ADC_h$ is still by far bigger than $ADC_r$. If the pore size distribution is fairly narrow then the system can be thought of as having two ADCs; $\mathrm{ADC}_h$ for the hindered bulk water, and a mean $\mathrm{ADC}_r$ for all pores.

At sufficiently long observation times, the ADC can be expressed as the weighted sum of two diffusion coefficients, $ADC_h$ for the time during which the particles were outside the walls and $ADC_r$ for the times when the particles were inside the walls.

$$ADC = (1 - P_i) \times ADC_h + P_i \times ADC_r \tag{2.43}$$

where $P_i$ is the expected fraction of time spent inside the pores. It is equal to the total volume fraction $\Phi$ occupied by all the small pores. Combing Eq. 2.42 with Eq. 2.43 gives

$$ADC = (1 - \Phi) \times ADC_h + \Phi \times ADC_r \tag{2.44}$$

We need to note here that the ADCs discussed above are only related to the mean squared displacement and do not imply Gaussianity of the diffusion propagator. The ADC is only the second order approximation of the propagator - the propagator itself is better approximated by two Gaussians.

## 2.7 Discussion

This chapter introduced the basics of diffusion (i) in unrestricted, isotropic form, (ii) diffusion as a tensor, and (iii) the diffusion propagator. The diffusion propagator was described as the probability of a particle starting at $\vec{r}_0$ and ending at $\vec{r}$ at time $t$ and it is sufficient to describe diffusion motion in all forms with or without the presence of geometrical obstructions or local inhomogeneities. The diffusion propagator was elaborated for simple restricted geometries that are widely used as model setup in literature and will be referred to in subsequent chapters. The ensemble-averaged propagator (EAP) that is measured with nuclear magnetic resonance was explained.

It is important to note that though EAP, in principle, could provide the shape as well as dimensions of geometrical restrictions, it involves many parameters: scaling a few parameters could produce similar result from two different geometries.

The final section was a discussion on some ways to understand diffusion in heterogeneous system where complete analytical descriptions are not possible. Most diffusion studies at present are focused on these disordered systems where both the mathematical formulation as well as physical understanding is still limited.

## 2.8 Appendix

Recurrence relation for Bessel functions

$$J_{n-1}(x) = \frac{n}{x} J_n(x) + J_n^{'}(x) \tag{2.45}$$

$$J_{n+1}(x) = \frac{n}{x} J_n(x) - J_n^{'}(x) \tag{2.46}$$

Lommel's Integral for Bessel Functions:

$$\int_0^1 r J_n(\alpha r) J_n(\beta r) dr = \begin{cases} \frac{1}{\beta^2 - \alpha^2} \left[\alpha \; J_n{}'(\alpha) J_n(\beta) - \beta \; J_n(\alpha) J_n{}'(\beta)\right] \;, & \alpha \neq \beta \\ \frac{1}{2} \left[ \; J_n^2(\alpha) - \; J_{n-1}(\alpha) J_{n+1}(\alpha)\right] & \alpha = \beta \end{cases} \tag{2.47}$$

$$J_n(x) = \frac{1}{\pi} \int_0^{\pi} \cos(n\tau - x \sin(\tau))\, d\tau \tag{2.48}$$

$$J_n(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-i(n\tau - x \sin(\tau))}\, d\tau \tag{2.49}$$

$$J_n(x) = \frac{i^{-n}}{\pi} \int_0^{\pi} e^{ix \cos(\tau)} \cos(n\tau)\, d\tau \tag{2.50}$$

$$(\beta^2 - \alpha^2) \int x J_n(\alpha x) J_n(\beta x)\, dx = x \left\{ \alpha J_n^{'}(\alpha x) J_n(\beta x) - \beta J_n^{'}(\beta x) J_n(\alpha x) \right\} \tag{2.51}$$

# Chapter 3

# Nuclear Magnetic Resonance

Nuclear magnetic resonance (NMR) principles rely on the existence of a nuclear angular momentum, also known as spin, and its interaction with the external magnetic field. Since the discovery of spin states with Stern-Gerlach experiment 90 years ago, most of our understanding on nuclear spin has come through its interaction with the external magnetic field.

NMR is also one of the best known examples of a quantum mechanical phenomenon that has found a widespread use. The discovery of spins has allowed us to understand molecular environments with minimal experimental perturbation. To name a few, NMR allows to observe precession frequencies, relaxation rates, couplings and translational motion.

Although spin and its interactions do not have counterparts in classical physics, at temperatures where Boltzmann statistics prevails, the macroscopic behavior of an ensemble of spins can be described with principles of electromagnetism. In this chapter, I will briefly describes few basics of quantum mechanics required for an overview of the resonance phenomenon. We will describe a spin, its mathematical representation and how it manifests in a measurement. The Zeeman interaction has been described with quantum mechanical laws followed by a move towards classical Bio-Savart law to describe induction and detection of an electromotive force. The Bloch equations, spin echoes and measurement of translational motion have been described as needed for later chapters. The final section describes use of NMR to understand diffusion of spin carrying molecules.

## 3.1 Quantum States and Nuclear Magnetism

Nuclear spin is a fundamental property that is described by the nuclear angular momentum quantum number, $I$. The nucleus of an atom is many times heavier than the electron revolving around it. The orbital angular moment for the nucleus is virtually

zero and the spin quantum number is equivalent to its angular moment quantum number. However, spin is not equivalent to actual motion. In this regard spin is a truly quantum mechanical property of an elementary particle.

### 3.1.1 Quantum states

Measurement of the spin $I$ results in the projection of one of its components along the observation axis, $z$.

$$I_z \left|m\right\rangle = m \left|m\right\rangle \tag{3.1}$$

Here $I_z$ is the operator for angular momentum in $z$-axis and $m$ is the eigenvalue corresponding to the observational possibility of each state $|m\rangle$. These observational possibilities are also called azimuthal quantum number $m$. The measurement results only discrete integer or half-integer values between $-I$ and $I$.

A nuclear spin system with angular momentum $I$ can be described in terms of its basis states

$$\left|\Psi\right\rangle = \sum_m a_m \left|m\right\rangle \tag{3.2}$$

where $a_m$ are the amplitudes of each basis state $m$. It contains both magnitude and phase such that $a_m = |a_m| \; exp(\, i\delta_\alpha \,)$.

For a nuclear spin in state $|\Psi\rangle$, measurement of any observable $A$ can be described as $\langle\Psi| \, A \, |\Psi\rangle$. Measurement of angular momentum in z-axis is

$$\begin{aligned} \left\langle\Psi\right| I_z \left|\Psi\right\rangle &= \sum_{m,m'} a_m^* a'_m \left\langle m' \right| \, I_z \left|m\right\rangle \\ &= \sum_{m,m'} a_m^* a'_m \, m \left\langle m'|m\right\rangle \\ &= \sum_m \left|a_m\right|^2 m \end{aligned} \tag{3.3}$$

The algebra of the angular momentum follows commutation relationship, a full description is beyond the scope of this thesis and can be found in reference [19]. The discussion here is limited to a proton and its interaction with the magnetic field.

The proton is spin $\frac{1}{2}$ particle which can only have two possible values of $\pm\frac{1}{2}$ for $I_z$. For a two level system like a proton, the basis of Pauli matrices $\{\mathbf{1}, \sigma_x, \sigma_y, \sigma_z\}$ greatly simplify the mathematics of the spin dynamics. Here, $\sigma = \frac{1}{2}(\mathbf{1} + \hat{n} \cdot \vec{\sigma})$ and $\hat{n} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$. For a spin-1/2 system, the vector $\hat{n}$ characterizes the polarization of the spin. The Pauli matrices for a spin 1/2 can be written as

$$I_x = \frac{1}{2}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, I_y = \frac{1}{2}\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, I_z = \frac{1}{2}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{3.4}$$

and **1** is the identity matrix. The basis vectors for spin up and down are

$$\left|\frac{1}{2}\right\rangle = \begin{bmatrix}1\\0\end{bmatrix} \; and \; \left|-\frac{1}{2}\right\rangle = \begin{bmatrix}0\\1\end{bmatrix} \tag{3.5}$$

The use of Pauli matrices will be seen in the following section.

### 3.1.2 Nuclear Magnetism

A nuclear magnetic dipole moment, $\vec{\mu}$, is proportional to the spin, $I$,

$$\vec{\mu} \; = \; \gamma I \tag{3.6}$$

where $\gamma$ is the constant of proportionality known as gyromagnetic ratio.

Named after Zeeman, the interaction between a magnetic dipole, $\vec{\mu}$, and the external field, $\vec{\mathbf{B}}_0$ is called Zeeman interaction. It is expressed as $-\vec{\mu}\,\vec{\mathbf{B}}_0$. The difference in Zeeman interaction energy between the two spin states of a proton is $\gamma\hbar B_0$. Similar to photoelectric effect, transition from one state to another is facilitated by absorption of emission of a photon with energy $\hbar\omega = E = \gamma\hbar B_0$.

Evolution of the nuclear spin with time can be described with Schrödinger equation.

$$i\hbar\frac{\partial}{\partial t}\left|\Psi(t)\right\rangle \; = \; H\left|\Psi(t)\right\rangle \tag{3.7}$$

where $H$ is the Hamiltonian. As in modern NMR spectrometers, if the external magnetic field is constant and the magnetic field $B_0$ is oriented along $z$-axis,the Hamiltonian, $H$, can be expressed as $H \; = \; -\gamma\hbar B_0 I_z$, and

$$\left|\Psi(t)\right\rangle \; = \; exp\,(-\frac{iHt}{\hbar})\left|\Psi(0)\right\rangle \; = \; U(t)\left|\Psi(0)\right\rangle \tag{3.8}$$

is the solution to Eq. 3.7. The initial factor $U(t)$ is the evolution operator. It is equivalent to

$$U(t) \; = \; exp\,(-\frac{iHt}{\hbar}) \; = \; exp\,(i\gamma B_0 I_z t) \; = \; exp\,(i\omega t I_z) \tag{3.9}$$

where $\omega = \gamma B_0$ is called the Larmor frequency. The evolution operator has the same mathematical form as rotation along z-axis by an angle $\omega I_z t$. It appears as if an external magnetic field obliges all spins to precess at frequency $\omega = \gamma B_0$. When $B_0$ is not aligned

with $I_z$, the Hamiltonian can be easily expressed using Pauli matrices. In this case,

$$\begin{aligned} H &= -\gamma \left(B_x I_x + B_y I_y + B_z I_z\right) \\ &= -\gamma \left(\frac{1}{2}B_x \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + \frac{1}{2}B_y \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} + \frac{1}{2}B_z \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\right) \\ &= -\frac{\gamma}{2} \begin{bmatrix} B_z & B_x - iB_y \\ B_x + iB_y & -B_z \end{bmatrix} \end{aligned} \tag{3.10}$$

Using this formalism, the Schrödinger equation 3.7 for a proton can be easily rewritten as

$$i\hbar \frac{\partial}{\partial t} \begin{bmatrix} p_{[+\frac{1}{2}]} \\ p_{[-\frac{1}{2}]} \end{bmatrix} = -\frac{\gamma}{2} \begin{bmatrix} B_z & B_x - iB_y \\ B_x + iB_y & -B_z \end{bmatrix} \begin{bmatrix} p_{[+\frac{1}{2}]} \\ p_{[-\frac{1}{2}]} \end{bmatrix} \tag{3.11}$$

When the magnetic field is parallel to the z-axis, $B_x$ and $B_y$ are zeros and the equation reduces to

$$\begin{bmatrix} p_{[+\frac{1}{2}]} \\ p_{[-\frac{1}{2}]} \end{bmatrix} = \begin{bmatrix} A_{[+\frac{1}{2}]} exp(\frac{i\gamma B_z}{2\hbar}) \\ A_{[-\frac{1}{2}]} exp(\frac{-i\gamma B_z}{2\hbar}) \end{bmatrix} \tag{3.12}$$

Here $A_{[\pm\frac{1}{2}]} = |p_{[\pm\frac{1}{2}]}|^2$ are the magnitudes of the probabilities of the quantum states and the exponential terms give the evolution operator $U(t)$. This equivalent description helps describing spin dynamics from a classical viewpoint [20].

### 3.1.3 Probabilities of quantum states

For large number of spins, even though the state of an individual spin is not known, the ensemble can be described in terms of the occurrence probability for each spin state. At room temperature and pressure, relative frequencies of each spin state follows Boltzmann statistics. The probability of a spin occupying a given state $m$ is real number and is equivalent to squared amplitude of each state

$$|p_m|^2 = \frac{exp(-\gamma B_0 m\hbar/k_b T)}{\sum_m exp(-\gamma B_0 m\hbar/k_b T)} \tag{3.13}$$

For a proton immersed in a constant magnetic field $B_0$, $m$ takes only two possible values. In this case,

$$|p_{[\pm\frac{1}{2}]}|^2 = \frac{exp(\pm\hbar\gamma B_0/2k_B t)}{exp(\hbar\gamma B_0/2k_B t) + exp(-\hbar\gamma B_0/2k_B t)} \tag{3.14}$$

At room temperature and pressure $\hbar\gamma B_0 \ll 2k_B t$. Taylor expansion of Eq. 3.14 gives

$$|p_{[\pm\frac{1}{2}]}|^2 \cong \frac{1}{2}(1 \pm \frac{\hbar\gamma B_0}{2k_B T}) \tag{3.15}$$

and,

$$|p_{[+\frac{1}{2}]}|^2 \; - \; |p_{[-\frac{1}{2}]}|^2 \; = \; \frac{1}{2}\left(\frac{\hbar\gamma B_0}{k_B T}\right) \tag{3.16}$$

The difference between probabilities of the two states increases with increasing magnetic field. This difference in the relative population of the spin states is called polarization. Greater polarization at higher field consequently allows for a higher intrinsic SNR during detection.

In thermal equilibrium, the ensemble can be described as

$$|\Psi\rangle \; = \; p_{[+\frac{1}{2}]} \left|\frac{1}{2}\right\rangle + p_{[-\frac{1}{2}]} \left|-\frac{1}{2}\right\rangle \tag{3.17}$$

For this ensemble, following Eq. 3.3, the measured angular momentum $I_z$ can be written as

$$\begin{aligned} \overline{\langle I_z \rangle} \; &= \; \overline{|p_{[-\frac{1}{2}]}|^2 \left\langle \frac{1}{2} \right| I_z \left| \frac{1}{2} \right\rangle + |p_{[-\frac{1}{2}]}|^2 \left\langle -\frac{1}{2} \right| I_z \left| -\frac{1}{2} \right\rangle} \\ &= \; \frac{1}{2}\left(\overline{|p_{[\frac{1}{2}]}|^2} - \overline{|p_{[-\frac{1}{2}]}|^2}\right) \end{aligned} \tag{3.18}$$

Similarly, $\overline{\langle I_x \rangle}$ and $\overline{\langle I_y \rangle}$ can be written as

$$\overline{\langle I_x \rangle} \; = \; \frac{1}{2}\left(\overline{p^*_{[\frac{1}{2}]} p_{[-\frac{1}{2}]}} + \overline{p_{[\frac{1}{2}]} p^*_{[-\frac{1}{2}]}}\right) \tag{3.19}$$

$$\overline{\langle I_y \rangle} \; = \; -\frac{i}{2}\left(\overline{p^*_{[\frac{1}{2}]} p_{[-\frac{1}{2}]}} - \overline{p_{[\frac{1}{2}]} p^*_{[-\frac{1}{2}]}}\right) \tag{3.20}$$

While Eq. 3.18 describes spin polarization Eq. 3.19 and 3.18 relate to phase coherence. The polarization and coherence can be conveniently described with density matrix.

$$\rho = \begin{bmatrix} \frac{1}{2} + \overline{\langle I_z \rangle} & \overline{\langle I_x - iI_y \rangle} \\ \overline{\langle I_x + iI_y \rangle} & \frac{1}{2} - \overline{\langle I_z \rangle} \end{bmatrix} = \begin{bmatrix} \overline{|p_{[\frac{1}{2}]}|^2} & \overline{p_{[\frac{1}{2}]} p^*_{[-\frac{1}{2}]}} \\ \overline{p^*_{[\frac{1}{2}]} p_{[-\frac{1}{2}]}} & \overline{|p_{[-\frac{1}{2}]}|^2} \end{bmatrix} \tag{3.21}$$

The density matrix contains all the necessary information for the ensemble and makes it convenient to express observations on an ensemble. The diagonal elements describe the population density of the two states. The off diagonal elements describe the coherence between the phase associated with these states. The off-diagonal terms are called the transverse spin magnetization. Coherence also requires the two transverse magnetization to be aligned. The Larmor precession does not affect the density matrix.

The density matrix greatly helps with the mathematics of the time evolution of the overall spins. The time evolution can be written as

$$\rho(t) = U(t)\rho(0)U^{-1}(t) \tag{3.22}$$

Eq. 3.22 is known as Liouvelle equation. There use will be seen in the following section.

## 3.2 Excitation and Signal Detection

The precessing spins can be perturbed with by an electromagnetic field. A radio-frequency (RF) pulse tuned at the Larmor frequency acts as an oscillating magnetic field. The quantum mechanics behind spin transition from one state to another was worked out by Rabi [20].

### 3.2.1 Excitation

In the presence of an oscillating magnetic field perpendicular to the static magnetic, the dynamics of the spin system can be described by referring back to Eq. 3.11 and adding the oscillating terms. Let us consider a small RF field $B_1$ which is applied perpendicular to $B_0$ and rotating around with but with a far smaller magnitude $|B_1| << |B_0|$. The $B_1$ field can be described by

$$B_x = B_1 cos(\omega t), B_y = B_1 sin(\omega t) \tag{3.23}$$

Using Pauli matrices (Eq. 3.10), the Hamiltonian can be written as

$$H = -\frac{1}{2}\begin{bmatrix} \omega_0 & \omega_1 e^{-i\omega t} \\ \omega_1 e^{i\omega t} & -\omega_0 \end{bmatrix} \tag{3.24}$$

where $\omega_1 = \gamma B_1$. In the presence of the RF field, the time evolution can be expressed similar to Eq. 3.11. However, the calculation is much easier where we look at this time-dependent Hamiltonian in the rotating frame of reference. In the coordinate frame, that is rotating with the Larmor frequency, the Hamiltonian can be written as

$$H = -\frac{1}{2}\begin{bmatrix} \omega_0 - \omega & \omega_1 \\ \omega_1 & \omega - \omega_0 \end{bmatrix} \tag{3.25}$$

In this sense $\omega$ is the frequency of nutation. It is also known as the Rabi frequency.

The RF pulse acts as an rotation operator, $U^{nut}(t_p) = exp(i\omega t_p I_x)$.

$$U^{nut}(t_p) = \begin{bmatrix} cos(\frac{1}{2}\omega_1 t_p) & isin(\frac{1}{2}\omega_1 t_p) \\ i\, sin(\frac{1}{2}\omega_1 t_p) & cos(\frac{1}{2}\omega_1 t_p) \end{bmatrix} \tag{3.26}$$

Using Eq. 3.22, the evolution of the density matrix, $\rho$ after the nutation operator can be described as

$$\rho(t+) = U^{nut}(t_p)\rho(t-)\left(U^{nut}(t_p)\right)^{-1} \tag{3.27}$$

where $\rho(t+)$ is the final density matrix after the RF pulse and $\rho(t-)$ is the initial density matrix before the application of the RF pulse. Before, the RF pulse, the spins are in thermal equilibrium and as such $\rho(t-)$ is

$$\rho(t-) = \begin{bmatrix} p_{[\frac{1}{2}]} & 0 \\ 0 & p_{[-\frac{1}{2}]} \end{bmatrix} \tag{3.28}$$

When $\omega_1 t_p$ is exactly equal to $\pi/2$, the final density matrix $\rho(t+)$ turns out to be

$$\rho(t+) = = \frac{1}{2} \begin{bmatrix} 1 & -i(p_{[\frac{1}{2}]} - p_{-[\frac{1}{2}]}) \\ i(p_{[\frac{1}{2}]} - p_{-[\frac{1}{2}]}) & 1 \end{bmatrix} \tag{3.29}$$

The RF pulse that places the spins on the transverse plane is called a $\pi/2$ or a 90°. They produce the maximum signal during detection. RF pulses can be applied at any arbitrary angle and multiple RF pulses are also possible. An important RF pulse is the 180° which acts as a reflector of the spin magnetization: it switches the polarization state and flips the sign of the phase coherence.

The coherence in the transverse plane is detected as NMR signal.

### 3.2.2 Signal Detection

When the magnetization,$\vec{M}$, is in the transverse plane, the Larmor evolution operator acts on the density matrix whose state changes at the rate of omega. Although nuclear magnetic momentum is a quantum mechanical quantity, any two-level system can be written in the form $e^{-i\phi/2}\cos(\theta/2)|0\rangle + \sin\theta(\theta/2)e^{i\phi/2}|1\rangle$ justifying the Bloch sphere interpretation. As Bloch points out in his seminal paper [21]

*"This variation does not require solution to the Schrödinger equation. It is enough to remember the general fact that the quantum mechanical expectation value of any quantity follows in its dependence exactly the classical equations of motion and that the magnetic and angular momenta of each nucleus are parallel to each other."*

Analogous to Eq. 3.6 one can describe the net magnetic moment, $\vec{M} = \gamma A$, where $A$ is the net angular momentum of all vectors. On the other hand, the torque $T$ can be described in terms of the angular momentum as

$$\frac{dA}{dt} = T = \vec{M} \times \vec{B} \tag{3.30}$$

which then follows

$$\frac{d\vec{M}}{dt} = \gamma \vec{M} \times \vec{B} \tag{3.31}$$

These equations are similar in form to Maxwell's equations.

In classical terms we can think of Faraday induction: put an antenna and detect this rate of change of magnetization. This signal detection through induction is termed as free induction decay (FID). As the magnetization $M(t)$ rotates around $B_0$ at the Larmor frequency, the induced voltage at the receiver also oscillates.

In modern scanners, Larmor frequency ranges in hundreds of MHz and detection of such a signal would be heavily influenced by cosmic noise. Although the oscillations are at such high frequencies, the frequencies of interest are within a limited bandwidth. This allows for frequency demodulation of the signal. In simple terms demodulation is achieved by multiplying the incoming signal with another oscillating voltage with a frequency close to the Larmor frequency. Mathematically the multiplication results in two waveforms, whose frequencies are sum and difference between the two input frequencies. The electronics time-averages the higher frequency retaining a limited bandwidth near zero. In a way, demodulation puts the incoming signal into the rotating frame and allows us to time-average the signal during acquisition.

If the reference waveform used for demodulation is shifted by quarter of a period, one can gain another component of the signal that is orthogonal to the first one. Only when the Larmor frequency and the reference frequency are equal and on-phase, the time average of this component is zero. For all other cases, this orthogonal component provides us with extra information. This process of demodulating the initial signal with two waveform is called heterodyne mixing. Here, the initial induced signal is split into two, and each half is mixed with two reference waveforms with same frequency but phase-shifted by a quarter wave. Mixing the signal in this way results in projection of the signal is two orthogonal axis. The final signal is then expressed as complex number. Heterodyne mixing helps during Fourier transform of the signal and thus forms and integral part of what is known as Fourier transform NMR. As we will later see, the Fourier transform of the signal forms the basis of most NMR theory and applications.

## 3.3 Spin Dance between Excitation and Signal Detection

The magnetization immediately after excitation is determined by the excitation pulse. With the passage of time the remaining magnetization and coherence is determined by different factors that are both external and internal to the system. The final signal reflects the net effect of these factors on the ensemble of the spins. This section briefly describes the major effects.

### 3.3.1 Spin-lattice Relaxation

Once the magnetization is perturbed from equilibrium, the spin polarization moves back towards equilibrium while releasing energy to its surrounding. While the individual spins randomly flip up and down, the net magnetization slowly moves towards equilibrium. If

the response of each individual spin carrying particle back is independent and randomly occurring, then the net change in spin density within a given time $dt$ is proportional to the difference between net polarization at that time and the polarization at the equilibrium. The net change in polarization can be expressed as

$$\frac{dM_z}{dt} = \frac{M_0 - M_z}{T_1} \tag{3.32}$$

where $M_z$ is the net polarization at time $t$ and $M_0$ is polarization at equilibrium. $T_1$ is the proportionality constant known as spin-lattice relaxation time.

### 3.3.2 Phase Distribution and Signal

While $T_1$ tells us about the time taken by the spins to get back to their equilibrium polarization after an excitation pulse, the NMR signal itself also depends on the coherence of the spins precessing in the transverse plane. In an ideal situation when all the spins are completely non-interacting and each of them precess at the Larmor frequency, then due to heterodyne, we would find the real part of the signal slowly decaying with $T_1$ and zero imaginary part. However, in many situations each individual spin could deviate from Larmor precession and gain an effective phase with the passage of time. The signal then is the complex average of the signal from each spin. The normalized signal is

$$S(t) = \langle e^{i\Delta\phi} \rangle = \int_{-\infty}^{\infty} P(\Delta\phi) e^{i\Delta\phi} d(\Delta\phi) \tag{3.33}$$

$e^{i\Delta\phi}$ is the Fourier transform of the P. The Fourier relationship between phase and the signal is foundational in understanding the NMR signal behavior.

### 3.3.3 Spin-spin Relaxation

In the molecular scale, spins dipole-dipole interactions of the molecules or chemical shielding due to a particular electronic could cause variations in the Larmor frequency at different times. With the passage of time, each spin gains some phase, $\Delta\phi$, that reflects its cumulative deviation from the Larmor frequency. For fluids, the time scale of these variations are on the order of picoseconds: by virtue of central limit theorem the distribution $P(\Delta\phi)$ is a normalized Gaussian whose variance increases linearly with time $t$. For a Gaussian PDF, the Fourier transform can be conveniently expressed.

$$\langle e^{i\Delta\phi} \rangle = e^{-\frac{\langle \Delta\phi^2 \rangle}{2}} \tag{3.34}$$

where $\langle \Delta\phi^2 \rangle$ is the variance of the PDF which increases linearly with time $t$. Thus, we can write the signal as an exponential function of time.

$$\langle e^{i\Delta\phi} \rangle = e^{-R_2 t} \tag{3.35}$$

where $R_2$ is the constant of proportionality and its inverse $T_2 = 1/R_2$ is known as spin-spin relaxation time. Its corollary

$$\frac{dM_{x,y}}{dt} = -\frac{M_{x,y}}{T_2} \tag{3.36}$$

is the Bloch equation for spin-spin relaxation. Since longitudinal relaxation would also result in a loss of spin coherence, $T_2 \leq T_1$.

The FID obtained is usually Fourier transformed. The Fourier transform of an exponential decay is a Lorentzian. The integral of this Lorentzian is related to the number of spins in our sample and its bandwidth reveals the $T_2$.

### 3.3.4 Relaxation due to gross inhomogeneities

While spin-spin relaxation is intrinsic to spin carrying molecules, heterogeneity could also exist on a bigger scale due to inhomogeneities in the Larmor frequency. Considering linear field gradient, for such 'inhomogeneous broadening' of the Larmor frequency, the Bloch equation can be expanded as

$$\frac{\partial}{\partial t} M_{x,y} = -i\gamma \vec{r}.\vec{g} - \frac{M_{x,y}}{T_2} \tag{3.37}$$

Here, $\vec{r}$ is the position of each spin and $\vec{g}$ is the gradient of the magnetic field. Depending on their position each spin rotates with slight deviation from the Larmor frequency so that the phase coherence is rapidly lost.

Fortunately, unlike $T_2$ decay, this loss in phase coherence is reversible. Hahn [22] showed that the first term in the above equation could be nulled by applying a 180° pulse refocusing pulse at a time $t$. The 180° pulse inverts the phase of each spin so that at time $TE = 2t$ all spins are once again in phase with each other and a spin echo is formed. Without further digression, it is noted that the 180° pulse can be applied along any axis in the transverse plane. Application of an multiple RF pulses after a single excitation must be carefully assessed for relative phases between the pulses and multiple spin echoes that would subsequently arise. In Hahn's honor, spin echoes are also called Hahn echo.

A related method is to use 90° RF pulse instead of the 180° pulse. When a 90° pulse is applied immediately after the first excitation pulse, so that spins do not dephase, then the two pulse together can act as one 180° pulse. However, when the second 90° pulse is applied after the spins have dephased, it sends part of the magnetization in the

longitudinal axis, while keeping part of it in the transverse plane. The ratio between magnetization in longitudinal axis and in the transverse plane depends on the dephasing. If the spins have been completely dephased before the first and the second pulse, then exactly one half of the effective magnetization is sent to the longitudinal axis. The other half remaining at the transverse plane still produces Hahn echo but the one in the longitudinal axis preserves it phase. The longitudinal magnetization only undergoes $T_1$ relaxation and can be later recalled by applying another 90°. This mode of storing magnetization in the z-axis can be used to measure $T_1$ relaxation, but since $T_1$ relaxation is always longer that $T_2$, this method can also be used to measure diffusion at long time. The method of measuring diffusion with three 90° RF pulses is known as stimulated-echo acquisition mode (STEAM). STEAM will be again encountered in section 6.1.

Hahn also realized that the resulting spin echo is sensitive to difference in spin trajectories before and after the refocusing pulse [22]. Therefore, spin echoes are still sensitive to flow and diffusion of the spins.

Signal formation due to inhomogeneities was first theoretically developed by Carr and Purcell [23] and later Torrey modified Bloch equation to include spin diffusion and flow [24]. The equation famously knows as the Bloch-Torrey equation is

$$\frac{\partial}{\partial t}M_{x,y} \;=\; -i\gamma\vec{r}\vec{g}(t) - \frac{M_{x,y}}{T_2} + \nabla(\mathbf{D}(\vec{r})\nabla M_{x,y}) - \nabla\vec{v}M_{x,y} \tag{3.38}$$

where $\mathbf{D}$ is the diffusion coefficient of the spin ensemble and $\vec{v}$ is the velocity term. Stejskal and Tanner [25] showed that the solution to Bloch-Torrey equation is

$$M_{x,y} = S(t)e^{-i\langle\psi(t)\rangle}e^{(-t/T_2)} \tag{3.39}$$

where

$$\psi(t) \;=\; \gamma\int_0^t \vec{r}(t)\vec{g}(t')dt' \tag{3.40}$$

Adding a 180° pulse after time $t$ would result in a spin-echo at time $TE = 2t$ and

$$M_{x,y}(TE) = S(TE) \times exp(-TE/T_2) \tag{3.41}$$

where $S(TE)$ contains the effect due to the gradients.

$$\frac{\partial S(TE)}{\partial t} = -\left(\gamma\int_0^t(\vec{g})(t')dt'\right)D\left(\gamma\int_t^{TE}(\vec{g})(t')dt'\right) \tag{3.42}$$

The negative sign accounts for the change in phase due to 180° pulse. When the system is in steady state with no flow, the only contribution to $S$ is due to diffusion of the spins. Eq. 3.42 shows that to measure $T_2$ of a sample, diffusional effects in the spin echo can be minimized if the refocusing pulse is placed at shorter time intervals. This forms the

basis for Carl-Purcell's [23] technique to measure spin echoes. It includes a train of 180° pulse starting at time $\tau$ after excitation and then $2\tau$ subsequently thereafter.

Variations in the Larmor frequency is a nuisance for $T_2$ measurements, but it can be exploited for two purposes

(i) to probe spatial distribution of the spin properties in a given sample - MR imaging, and

(ii) to understand diffusion properties of the spins.

## 3.4 MR Imaging

A basic MR imaging experiment exploits spatially varying magnetic field for two major purposes: (i) to select a spatially localized region of interest and (ii) to obtain an image of the sample.

### 3.4.1 Spatial Localization

In the previous section we discussed the rotation of spin density states in the presence of a RF pulse tuned at the Larmor frequency. This was achieved because in the rotating frame on-resonant spins remain static [26]. However, when the RF pulse is not 'on resonance', the off-resonant spins will precess with a frequency $\Delta\omega = \gamma\Delta B_0$. If the frequency offset is big enough RF pulse does not affect the net magnetization of the system. Spatial localization is a way to selectively excite a specific region of the frequency spectrum. The frequency bandwidth of an RF pulse is inversely proportional to its duration.

The simplest way to select a well-defined region is to locate a rectangular region along the magnetic field gradient. A linear magnetic field gradient together with RF pulse when applied will rotate spins which are located in a plane through the object [27]. If the RF pulse contains a band of frequencies $\delta\omega_1$ and a central frequency $\omega_1$, it will selectively excite spins that are within this bandwidth. For small angle excitations Hoult et al. [26, 28] showed that slice profile can be approximated by the Fourier transform of amplitude modulated RF pulse. To obtain a rectangular slice selection it is desired to have a RF pulse with a sinc profile, however in practice ideal sinc pulses are not possible at finite times and numerical methods for optimized RF pulses have been widely used [29, 30].

### 3.4.2 Obtaining sample image

If the sample is subjected to magnetic field gradient, the spins that were otherwise in-phase are not anymore. Here again the normalized signal is the sum of all spins and their magnetization. In the most general form, when the gradients are systematically

applied in all three axes, ignoring inhomogeneity in $B_0$ and multiple-spin populations in the sample, we can describe the signal as

$$S(G_x t, G_y t, G_z t) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \rho(x,y,z) e^{i(G_x t x + G_y t y + G_z t z)}\, dxdydz \tag{3.43}$$

Usually $\frac{\gamma G t}{2\pi}$ is expressed as $\vec{k}$ and the equation is written as

$$S(\vec{k}) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \rho(\vec{r}) e^{2\pi i \vec{k}.\vec{r}}\, dr \tag{3.44}$$

The spin density and k-space are thus the conjugate space of each other. Since the spin density is real, $S(-\vec{k}) = S*(\vec{k})$, so ideally only half of the k-space can be sampled to obtain a full image. When imaging is applied along with slice selection, then in most cases $k_z$ is completely dropped and k-space is thought in two dimensions only.

While the k-space sampling can be applied in any fashion. Cartesian acquisition of the k-space gives in easily to discrete Fourier transformed. For time efficiency, it is also a common practice to acquire all of k-space after single excitation. The most popular k-space acquisition scheme is called echo-planar imaging(EPI) [31].

## 3.5 Diffusion NMR

### 3.5.1 Stejskal-Tanner Pulse Sequence

Eq. 3.42 linked the fundamental relationship between diffusion and reduction in the spin echo amplitude. However, although elegant, this formulation is not quite handy when it comes to actually measuring diffusion. A convenient way to measure diffusion was invented by Stejskal and Tanner [25]. They developed the technique of pulsed gradient spin echo (PGSE). In the Stejskal-Tanner PGSE technique, a pair time-constant magnetic field gradients is applied for a short duration $\delta$ before and after the refocusing 180° pulse.

Although diffusion measurements had been done before using constant field gradients [16, 32], the introduction of pulsed field gradient greatly simplifies the mathematics as well as the understanding of diffusion measurements. PGSE is still the workhorse for MR based diffusion measurements. In what follows we describe the phase evolution in a PGSE sequence.

Figure 3.1 shows the simplified schematic of a PGSE experiment. Immediately after the first pulsed gradient, the spins gain some phase depending on their position.

$$\phi(\vec{r}, \tau_1) = \gamma \Delta(B_0(\vec{r})).\tau_1 + \gamma(\vec{g}.\vec{r})\delta \tag{3.45}$$

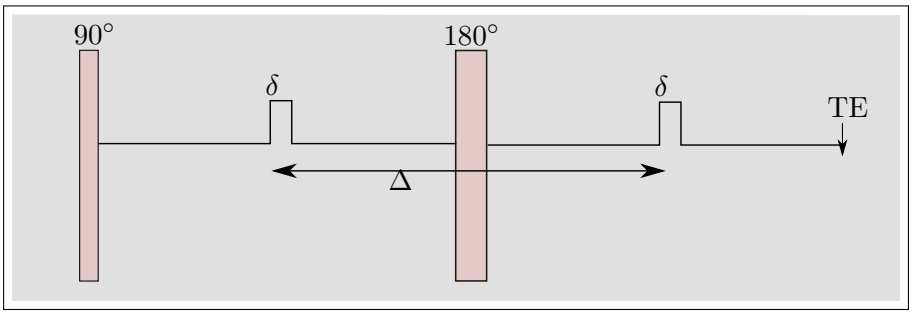



FIGURE 3.1: Schematic of a Stejskal-Tanner PGSE sequence

Here, $\vec{g}$ is the applied pulsed gradient. The first term is the inhomogeneous broadening (Eq. 3.37) and the second term is due to the pulsed gradient with duration $\delta$. If the inhomogeneity in the magnetic field is slowly varying then its effect is minimal at the Hahn echo. The second gradient pulse imparts an additional phase to the spins. At the Hahn echo, the phase can be written as

$$\phi(\vec{r}, TE) = \gamma(\vec{g}.\vec{r}(\tau_1))\delta - \gamma(\vec{g}.\vec{r}(\tau_2))\delta \tag{3.46}$$

The signal at Hahn echo is then

$$S(g, TE) = \int\int \rho(\vec{r_1}) P(\vec{r}_1|\vec{r}_2) e^{i\gamma\delta\vec{g}.(\vec{r}_1-\vec{r}_2)} d\vec{r}_1 d\vec{r}_2 \tag{3.47}$$

where $\rho$ is the density of spins, $\vec{r}_{1,2}$ are the positions of the spins at time $\tau_{1,2}$ and $P(\vec{r}_1|\vec{r}_2)$ is the probability associated with spin displacement from $\vec{r}_1$ to $\vec{r}_2$ during time $\tau_2 - \tau_1$.

At steady state the spatial density of spins over time is constant and thus NMR effectively measures the self-diffusion of the spins probed. Furthermore, when the spin density is constant over the sample, the signal is yet again the Fourier transform of the probability $P(\vec{r}_1|\vec{r}_2)$. In the first chapter, we described this probability as diffusion propagator. This Fourier relationship between the diffusion propagator and the gradient integral is commonly used in diffusion NMR and is known as q-space imaging. The term q-space is similar to k-space that we earlier encountered, here q-space is defined as $\gamma\delta g/2\pi$. We should note that the Fourier relationship is analytically valid for infinitely narrow gradient pulses.

### 3.5.2 Long Gradient Pulses, Gaussian Diffusion and B-value

It is not always practical to attain narrow gradient pulses. When gradients are applied for longer durations, dephasing due to spin motion during the pulse also plays a role.

To make the equation tractable we have conveniently ignored inhomogeneity in the magnetic field itself. After the first gradient 'pulse' (which is now significantly longer,

we call it a gradient lobe ) $\delta$, we find that the phase of a spin is

$$\phi(\vec{r}, \tau_1 + \delta) = \gamma(\vec{g}.\vec{r})\delta + \Phi(\delta) \tag{3.48}$$

Here, $\tau_1$ and $\vec{r}$ are time and position at the beginning of the gradient lobe and $\Phi(\delta)$) is the cumulative phase attained due to spin diffusion during this time. It is clear that $\Phi$ now depends not only on the initial position but also on the path of the spin during the gradient pulse. The spin-echo will nullify the phase accumulated due to initial position, and so it is enough for us to proceed as if the particles initially were at the origin ($\vec{r} = 0$).

For pure liquids, where the diffusive motion is truly random and no correlation exists between successive motion, the phase evolution is still tractable. In this case, the displacement $\vec{r}$ is always Gaussian and the instantaneous phase ($\phi = \vec{g}.\vec{r}$) is a continuous adiabatic operation on this Gaussian: therefore, due to central limit theorem $P(\Phi, \delta)$ is also a Gaussian. Nevertheless, Gaussianity of $\Phi$ lies at the heart of diffusion NMR and needs fuller attention.

An easy way to imagine this adiabatic process is to divide the gradient lobe $\delta$ into infinitely small d$t$ (Fig. 3.2). At each time interval d$t$, the spin gains a net displacement d$\vec{r}$. The phase due to this small displacement depends on the displacement and the gradient that subsequently follows. At some final time $t$, the total phase acquired due to d$\vec{r}$ and subsequent gradient profiles is added to the final phase of that spin.

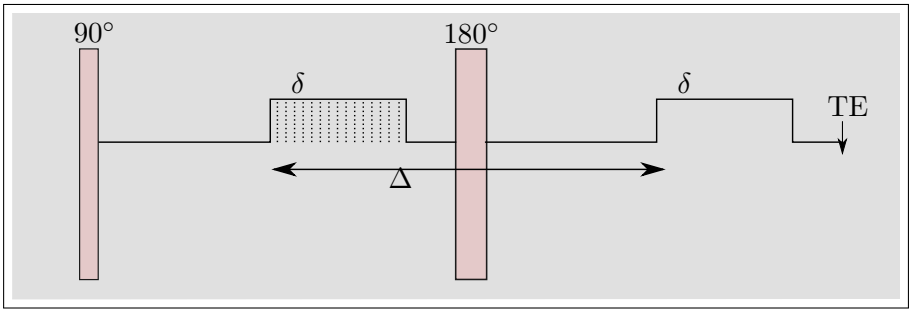



FIGURE 3.2: Schematic of a Stejskal-Tanner Pulse Sequence but with long gradient lobes. The gradients can be thought of as consisting of a many narrow gradients.

Thus the total phase accumulated due to d$\vec{r}(t')$ after some time $\tau$ is

$$\phi(t', \tau) = \gamma \int_{t'}^{\tau} (d\vec{r}(t').g(t))dt \tag{3.49}$$

The formulation can be seen as an impulse $d\vec{r}(t')$ at time $t'$ and the response $\phi(t', \tau)$ for this impulse. The PDF of the impulse does not necessarily need to be Gaussian, but its variance is well-defined, $\langle d\vec{r}(t')^2 \rangle = 6D$. In practice, $g(t)$ is applied in one direction. Diffusion is isotropic in bulk liquid so we consider $\vec{g}(t) = g_x(t)$, and the variance of $d\vec{x}(t')$ is then equal to $2D$.

We see that $\phi(t', \delta)$ are essentially identical distributions only multiplied by a factor $\int_{t'}^{\tau} g(t')dt$ depending on the gradient integral after time $t'$. Invoking central limit theorem, we can say that the total accumulated phase $\Phi(0, \tau)$ is also a Gaussian. The normalized signal is the Fourier transform of phase distribution (Eq. 3.33) and when the phase has a Gaussian probability distribution, the signal is directly related to the variance of phase (Eq. 3.34).

Since $\phi(t', \tau)$s are uncorrelated to each other, the variance of $\Phi(0, \tau)$ is the integral of all variances of $\phi(t', \tau)$. The variance of $\phi(t', \tau)$ can be written as

$$\begin{aligned} \langle \phi^2(t', \tau) \rangle &= \left[ \gamma \int_{t'}^{\tau} dx(t') g(t) dt \right]^2 \\ &= \gamma^2 \langle dx(t')^2 \rangle \int_{t'}^{\tau} g(t) dt \\ &= 2D\gamma^2 \int_{t'}^{\tau} g(t) dt \end{aligned} \tag{3.50}$$

For motion during the first gradient lobe,

$$\begin{aligned} \langle \Phi^2(0, \delta) \rangle &= \int_0^{\delta} \langle \phi^2(t, \delta) \rangle dt \\ &= 2D\gamma^2 \int_0^{\delta} \left\langle \left[ \int_{t'}^{\delta} g(t)\, dt \right]^2 \right\rangle dt' \\ &= 2D\gamma^2 g^2 \delta^3 / 3 \end{aligned} \tag{3.51}$$

The first line reminds of Eq. 3.42 from Bloch-Torrey formalism. To look at accumulated phase after the second gradient lobe we need to remember that the refocusing RF pulse flips the signs of the gradients. For simplicity, here we separate $\mathrm{d}\vec{r}(t)$ into three categories depending on if the motion occurred during the first gradient lobe, between first and second lobe or during the second lobe. Since all motions are uncorrelated, we can get variances for these three cases and simply add them to get the total variance. For motion that occurred during the first lobe

$$\begin{aligned} \langle \Phi^2(0 < t < \delta, TE) \rangle &= 2D\gamma^2 \int_0^{\delta} \left\langle \left[ \int_{t'}^{\delta} g(t) dt - \int_{\Delta}^{\Delta+\delta} g(t) dt \right]^2 \right\rangle dt' \\ &= 2D\gamma^2 g^2 \delta^3 / 3 \end{aligned} \tag{3.52}$$

Similarly for the motion that occurred between the first and the second lobe is

$$\begin{aligned}\langle\Phi^2(\delta < t < \Delta, TE) &= 2D\gamma^2 \int_{\delta}^{\Delta} \left\langle \left[\int_{\Delta}^{\Delta+\delta} -g(t)dt\right]^2 \right\rangle dt' \\ &= 2D\gamma^2 g^2 \delta^2 (\Delta - \delta) \end{aligned} \tag{3.53}$$

For the motion during the second gradient lobe, the phase accumulation is exactly as in Eq. 3.51. The variance in phase at time TE is thus, $2D\gamma^2 g^2 \delta^2 (\Delta - \delta/3)$.

The normalized signal is

$$S(TE) = e^{-D\gamma^2 g^2 \delta^2 (\Delta - \delta/3)} \tag{3.54}$$

The term $\gamma^2 g^2 \delta^2 (\Delta - \delta/3)$ is conveniently referred as the b-value.

B-value is a well-defined term to probe free bulk diffusion. It can be used to measure diffusion coefficient independent of the gradient profile. B-value can always be calculated by taking the time integral of the gradient square. B-value is proportional to the volume of the solid of revolution with time in x-axis and time integral of gradient on the y-axis.

## 3.6 NMR in heterogeneous medium

So far we have discussed behavior of NMR signal in homogeneous samples. However, NMR is perhaps the best measurement probe to understand physics of composite systems. In many cases it is the only tool that can be used to non-invasively observe system properties.

The easiest way to observe sample heterogeneity using NMR is with MR imaging (section 3.4). Imaging allows us to observe gross structure of the sample. Furthermore, even within a single imaging voxel, MR signal behavior reflects the underlying physics of the system that would otherwise not be visible at voxel level. However, for heterogeneous samples mathematical formulations are not so straightforward and interpretation of the NMR signal in heterogeneous medium is an active field of current research.

In this section we will briefly go through some simple cases for heterogeneous media and understand their effect in NMR signal evolution.

### 3.6.1 Non-mixing systems

The simplest case for heterogeneous system is when the system is composed of two non-interacting spin populations with distinct MR properties. An example of such a system is a mix of two different proton-containing molecules with no interaction between them. If the protons in the two molecules have different resonance frequencies then Fourier transform of the FID (NMR line shape) will show multiple peaks with different

bandwidth. Similarly if they have different $R_2$ relaxation rate then the $T_2$ decay will be a weighted sum of two exponential decays with their relative weights reflecting the density of the two spin populations. The diffusion decay curves will also be sum of two exponentials. The NMR line-shape also reflects inhomogeneity due to differences in macroscopic susceptibility. For most samples, however, the two spin populations interact with each other and the resulting NMR signal is not a simple sum. In the following discussion, we avoid any quantum mechanical interactions between the spins, like j-coupling, or polarization exchange and simply focus on exchange between different spin pools.

### 3.6.2 Two pool system with Exchange

#### 3.6.2.1 Spin Echo Signal

Here we consider SE signal from a sample with two spin pools, both precessing at the Larmor frequency but with distinct $R_2$ relaxation rate. We consider a case when exchange between the two pools is in the same order as $T_2$. This problem for relaxation with exchanging spins has been thoroughly investigated by Winkler et al. [33].

To solve this problem, we divide the final signal into four groups $S_{ij}$ where subscript $ij$ denotes the signal is contributed by spins that were in region $i$ at the beginning or our measurement and shifted to region $j$ at its end. Now up to first order,

$$\frac{dS_{ij}}{dt} = -R_j I_{ij} - \frac{S_{ij}}{\tau_j} + \Sigma_k \frac{S_{ik}}{\tau_k} P_{kj} \tag{3.55}$$

where $R_j$ is the relaxation rate in region $i$, $\tau$ is the mean time that the spin spends in each pool and $P_{kj}$ is the transition probability of a particle leaving region $k$ and going to region $j$.

For a system with just two regions, $P_{kj} = 1$. The equation can be expressed as a matrix.

$$\frac{d\mathbf{S}}{dt} = -\begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \times \begin{bmatrix} R_1 + 1/\tau_1 & -1/\tau_1 \\ -1/\tau_2 & R_2 + 1/\tau_2 \end{bmatrix} \tag{3.56}$$

Now, let's consider two apparent relaxation rates $R_1'$ and $R_2'$ that would give the same signal $S(t)$ as our exchanging spins with relative fractions $P_1'$ and $P_2'$. Then the apparent relaxation rates are eigenvalues of the second matrix in Eq. 3.56.

$$R'_{1,2} = \frac{1}{2} \times \left[ R_1 + 1/\tau_1 + R_2 + 1/\tau_2 \pm \sqrt{(R_1 + 1/\tau_1 - R_2 - 1/\tau_2)^2 + 4/(\tau_1\tau_2)} \right] \tag{3.57}$$

The corresponding volume fractions are

$$P_2' = \frac{R_1 P_1 + R_2 P_2 - R_1'}{R_2' - R_1'} \tag{3.58}$$

#### 3.6.2.2 Diffusion NMR signal

Similar to previous case, here we consider two spin pools with different intrinsic diffusivity, $D_1$ and $D_2$. Let's consider Stejskal-Tanner like diffusion sequence where the signal is observed at different TEs while keeping the diffusion gradients unchanged. With the same notations as before, Eq. 3.56 can be extended to include diffusion effects.

$$\frac{d\mathbf{S}(\mathbf{t})}{dt} = -\mathbf{S}(\mathbf{t}) \times \begin{bmatrix} R_1 + 1/\tau_1 + (\gamma\delta g)^2 D_1 & -1/\tau_1 \\ -1/\tau_2 & R_2 + 1/\tau_2 + (\gamma\delta g)^2 D_2 \end{bmatrix} \tag{3.59}$$

The equation above is commonly referred as Kärger equation in literature [34, 35]. From the equation above, it is immediately clear that the effect of diffusion and relaxation are additive. It is straightforward to first estimate relaxation rates by setting $g = 0$, and subsequently estimate the diffusion effect with some non-zero gradients.

### 3.6.3 Restricted System

In liquids, $T_2$ relaxation heavily relies on rapid rotational motion of spin-bearing molecules that effectively averages small differences in the Larmor frequency. For instance when water freezes, the Van der Walls force restricts rotational averaging and coherence is rapidly lost. Ice and most solids are therefore largely NMR invisible except for very short times.

Interaction with boundaries can also cause surface relaxation at the interface. The effect is more severe when boundaries are close enough to cause obstruction in molecular rotational motion. The water trapped in between lipid rafts are known to show reduced $T_2$. Similar argument is used to describe a short $T_2$ for water between myelin layers in brain WM.

However, for larger scale restrictions the relationship with $T_2$ is not clear. Let's consider spins residing in a confinement where rotational motion is not hindered but the space is small enough that the spins can diffuse and sample the space within a short time. It is clear that diffusive motion can average over tiny inhomogeneities due to field gradients at the surface, which would have otherwise caused loss of coherence. Of course, such effects due to the surface are not truly spin-spin relaxation and technically not $T_2$, but the difference is not always easy to distinguish experimentally.

Restrictions have a more direct relationship to diffusion. As discussed in the first chapter, diffusion depends largely on the size, geometry and permeability of the boundaries. The straightforward method to measure diffusion in simple restricting systems is with narrow gradient pulse.

The narrow pulse requirement assures that the diffusion length during the gradient pulse is significantly less that the pore size of the medium. In this case the signal is can

be written as

$$S(\gamma\delta g, \Delta) = \int\int \rho(\vec{r}_0) P(\vec{r}|\vec{r}_0, \Delta) \times e^{i\gamma\delta g.\vec{r}-\vec{r}_0} d\vec{r}_0 d\vec{r} \tag{3.60}$$

where $P(\vec{r}|\vec{r}_0, \Delta)$ is the diffusion propagator of the particle. A complete description of the signal in three geometries- parallel plates, cylinder and sphere, was calculated by Callaghan [36]. Using narrow pulses, diffusion-diffraction patterns have been reported for simple restricting geometries [37, 38]. Measuring such signal and inverting the Fourier transform can be used to reveal the average propagator.

However, in most situations applied gradients are long and phase evolution is less tractable. The narrow gradient pulse approximation breaks down but so does the Gaussian phase distribution. If the gradient duration $\delta$ is very long then diffusion during $\delta$ averages the phase resulting in a Gaussian phase PDF [39]. Similarly when gradients are weak, the phase PDF could be close to a Gaussian. This assumption of Gaussian phase is commonly referred to as Gaussian phase approximation (GPA). Assuming GPA, the signal is estimated from variance of the phase $\langle\Phi^2\rangle$. For diffusion in confined geometries the signal decay is found to be a function of the fourth power of the characteristic length of the restricting structure [40, 41]. The GPA approximated signal for the three simple geometries have been described by Neuman [39].

In the presence of sufficiently strong gradients, GPA does not hold anymore. The breakdown of GPA has been theoretically discussed by Swiet and Sen [42]. Hürlimann et al. [43] showed a stretched-exponential behavior for restricted diffusion between two parallel plates. The breakdown of GPA depends on the counter-effect between dephasing due field gradients and averaging of any such dephasing due to diffusion during the gradient pulse. The GPA breaks when the dephasing length $(D_0/\gamma g)^{1/3}$ is smaller than the diffusion length $\sqrt{D_0\delta}$ and the signal is not anymore exponentially decaying with b-factor.

## 3.7 Non-monoexponential diffusion decay

Non-monoexponential diffusion decay is encountered in many systems. It is common to see such behavior in most heterogeneous systems and is almost universally true in all biological systems. With the help of numerical simulations and analytical expressions, much effort has been directed on describing these decay phenomena and to extract relevant features of the sample. In the following section we will briefly describe two mathematical methods that have been employed in literature.

### 3.7.1 Cumulant Expansion

Though the GPA is no longer valid, the signal itself is still the Fourier transform of the phase distribution (Eq. 3.33). Recalling that for any bounded distribution $\Phi$, its Fourier

transform can be expressed in terms of its moments, the signal can be described in terms of the moments of $\Phi$ [44, 45].

$$S = \Sigma_{n=0}^{\infty} \frac{(iq)^n}{n!} \langle \phi^n \rangle \tag{3.61}$$

The equation above is the characteristic function whose Taylor expansion coefficients are raw moments of the phase PDF. The log of the signal can be expressed as a series representation of the cumulants of the distribution. Since the PDF is symmetric, odd moments are zero. The zeroth order moment is just a constant which can be normalized. The signal then is

$$ln(S) = \Sigma_{n=0}^{\infty} \frac{(-q^2)^n}{n!} \langle \phi^n \rangle_c \tag{3.62}$$

where $\langle \phi^n \rangle_c$ are cumulant moments of the phase distribution. When the phase is Gaussian, the fourth and higher order cumulant moments are zero. It is therefore the magnitude of these higher order cumulant moments that determine the validity of GPA. It should be noted that while Eq. 3.61 is convergent for all cases, Eq. 3.62 is only convergent close to the origin [46]. This limits the applicability of cumulant expansion methods for experiments with strong gradients.

In practice, the cumulant expansion method is simplified. The log of normalized signal is expressed in powers of $bD$ where $b$ is b-factor that was previously discussed with Gaussian diffusion and $D$ the diffusion coefficient. The series is usually truncated after the fourth order.

$$ln(S) = -bD + \frac{K}{6}(bD)^2 \tag{3.63}$$

where $K$ is the kurtosis that describes deviation of phase from Gaussianity.

### 3.7.2 Stretched Exponential

Stretched exponential is another way to mathematically expand the formalism. In this method, the normalized signal is expressed as

$$S = e^{-(bD)^{\alpha}} \rightarrow ln(S) = -(bD)^{\alpha} \tag{3.64}$$

where $0 < \alpha < 1$ is called the stretching factor. Smaller $\alpha$ usually correspond to higher heterogeneity in the system. Stretched exponential fit describes a special Gaussian mixture where the components are continuously distributed Gaussians with their volume fractions following a log-normal distribution.

## 3.8 Discussion

This chapter started from the basics of quantized spin states in protons, degeneracy due to magnetic field and interaction with an RF pulse. The interaction between spins and

surrounding environment is the basis of all NMR signal. The Fourier relation between observed signal and phase distribution was introduced and we saw how phases could be changed due to sample intrinsic factors like $T_2$ relaxation or external factors like varying magnetic field.

The mathematical foundation of MR imaging was discussed. It relies on the application of magnetic field gradients to obtain information on larger scale heterogeneities. Heterogeneities on smaller scales are still reflected on the phase behavior but needs more careful attention. Additionally, when the impact of two features ( like $T_2$ relaxation and diffusion) on the NMR signal is additive, multiple measurements can obtain information about both the properties.

While the NMR signal from Gaussian diffusion is relatively simple, it is not straightforward for a more complicated system. In such cases, the Gaussian phase approximation (GPA) comes handy and allows us to better approximate the expected signal behavior. When GPA is not valid, signal behavior can still be described by including higher moment of the phase PDF.

Since NMR signal is the ensemble average or all spins in a voxel, it is not a direct sum of contribution from all different regions. Especially for diffusion measurements, motion itself depends on the heterogeneity. We saw that the spins can easily exchange position between regions with different diffusion coefficients. For example, diffusion coefficients of fat-water mixture need not be equivalent to the diffusion coefficients of fat and water measured separately.

Biological tissues are extreme examples of such heterogeneous medium. Diffusion weighted signal from the tissue is not the same as signal from its components. Nevertheless, it does help to obtain information on the aggregate size, shape and composition of various structures that make up the system. In the following chapter we discuss the composition of brain WM and how different substructures would contribute to diffusion NMR signal.

# Chapter 4

# White Matter Architecture and Diffusion

*Fish can be found in murky water - Chinese Proverb*

The nervous system is responsible for gathering disparate information from different bodily systems, processing them as a holistic experience, and transmitting relevant signal back to the body. This complexity of functions is also reflected in its structure and composition. In developed animals, the nervous system can be further divided into peripheral nervous system (PNS) which is in contact with different bodily systems and central nervous system (CNS) which is responsible for processing and channeling information. The CNS, its structure and function, have been described in many volumes (See Ref. [47] and [48] for example). It consists of 3 major cells glial cell, oligodendrocytes and neurons and is broadly differentiated between gray matter/white matter (GM/WM) and brain/spinal cord. The brain WM is primarily responsible for crosstalk between different areas in the brain. The purpose of this chapter is to describe the constituents of brain WM and to address their role in water diffusion [1].

In WM the three basic building blocks are lipids, proteins and water. While water is the most abundant, occupying around 85% of volume in the GM and 75% of the volume in the WM, proteins and lipids organize in a breathtaking precision to form basic cellular substructures. Between different cell types, the substructures are similar with only small variations. The first part of the chapter briefly describes different WM substructures. Our approach is to describe WM like lego - larger structures are formed by smaller building blocks. Factors like micro-geometry, size and volume fractions and their influence on water diffusion will also be emphasized.

The second part of the chapter is a discussion on different diffusion models that have been used to describe diffusion measurements in white matter. The models are also

[1] From here onwards all references to WM is meant to be brain white matter unless explicitly stated otherwise.

considered in light of the structures described in the first half. The last part lays the foundation for diffusion measurements in the following chapters.

## 4.1 Microscopic Composites

### 4.1.1 Lipids

The lipid content of WM is reported to be around 49-66% of the dry weight [49] which translates to about 12-16% of the total WM mass. In the body, lipids are mostly found in three forms: as triglycerides in adipose tissue, as signaling molecules like hormones and as membrane building units. The adipose tissue is not present in WM. Hormones are involved in brain maturation and plasticity but in very small quantities. Most of the lipids in WM are present in membranes as phospholipids, sterols and glycolipids. All of these structures are amphipathic with a hydrophilic polar head and at least one hydrophobic hydrocarbon tail. It is energetically favorable for these lipids to have their polar heads close to water and the hydrophobic tails further away. Therefore, the lipids are mostly found as bilayers (See Fig. 4.1). The thickness of a lipid bilayer is approximately 10 nm.

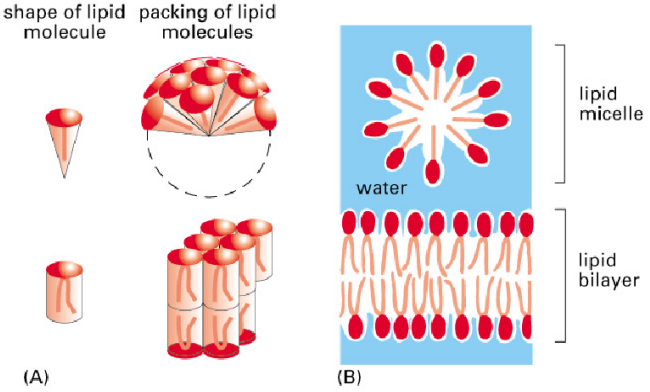



FIGURE 4.1: Different types of lipid structures.

Due to their hydrophobic nature, lipid bilayers act as solvent for lipophilic and amphipathic molecules and are permeable to small non-polar molecules like oxygen, nitrogen and carbon dioxide. But they are impermeable to ions and water soluble molecules. The water diffusion through the phospholipid bilayers is very low. The permeability of lipid bilayer is further reduced with increased cholesterol and sphingomyelin content [50].

Recent findings from terahertz time domain spectroscopy show that the phospholipid bilayers have a water hydration layer that extend up to 4-5 water layers [51]. These hydration water molecules are found to be ordered with the symmetry axis along the direction perpendicular to the bilayer [52]. Ordering reduces rotational as well as translational motion of water molecules, thus affecting their $T_2$ relaxation and ADC [53].

However, the exchange between hydration layer and neighboring free water is quite rapid (on the order of 10 picoseconds) and only a powder average can be observed by diffusion NMR measurements.

### 4.1.2 Proteins

More than 10% of the total WM by mass is composed of protein [54]. A typical cell is able to synthesize up to 20,000 different kinds of proteins and up to 1000 proteins have been identified in WM [55, 56]. Proteins are present in many forms in the WM with majority of them falling into three main categories: (i) as soluble proteins in the cytoplasm, (ii) as membrane protein embedded in the membranes and (iii) as cytoskeletal filaments.

#### 4.1.2.1 Soluble Proteins

Solubility of the cytoplasmic proteins depend on the difference in free energy between their solid state and when surrounded by water molecules. Here the folding conformation of the protein in aqueous solution plays an important role: a protein is soluble only if all of the exposed protein surface is made of polar groups that can interact with water molecules. Unlike polar solutions, protein-water are mostly colloids.

The soluble proteins are known to have solubility as high as 35%. A gram of fully hydrated protein interacts with about 0.35 grams of surrounding interfacial water molecules. The hydration layer of the interfacial water near protein is known to have different viscosity as well as rotational and translational properties than bulk water [57–59]. However the interface effects are short-ranged and residence time of water at protein surface is less than 1 ns [60]. Since diffusion-MR measurement probes water motion in millisecond time-scale, the protein-water colloid can be expected to give a reduced ADC due to both hydration layers and physical obstructions.

#### 4.1.2.2 Membrane Proteins

Membrane proteins are embedded within the rafts of lipid bilayers and typically contribute to about 50% of the membrane mass. While some proteins are simply associated with the membranes, other transmembrane proteins that span the lipid bilayer are important for through-membrane matter transport. Porin, for instance, is one such protein found in mitochondria that allows small molecules to pass in and out of the membrane.

Many membrane proteins act as ion channels. They are functionally as well as spatially selective. Sodium channels, for example, are initially found all along the axons but selectively congregate near the nodes of Ranvier after myelination [61]. Sodium channels are known to allow sodium ions but not potassium ions.

The ions channels are specific but also allow passage of hydrated protons associated with the ions. In this sense, these channels are also permeable to water molecules

and play a role in water diffusion. A specific kind of channels, called aquaporins, are specifically permeable to water molecules [62].

Channels can facilitate water diffusion across the membranes, but membrane proteins also hinder water motion along the membrane. Especially when the membrane bound spaces are small, proteins block most of the pathway.

#### 4.1.2.3 Cytoskeletal Filaments

The cytoskeletal filaments are complex structures made up of multiple protein components. As their name suggests, the cytoskeletal filaments act as scaffolds offering rigidity to the cells. Additionally, they are also involved in transporting macromolecules and cell organelles. The structural arrangement of cytoskeletal protein ensures their insolubility in water: individual proteins may be found in soluble state. For example, tubulin, the subunit of microtubules, is water soluble in its monomer state.

The cytoskeletal filaments are categorized into three: (i) microfilaments (ii) neurofilaments and (iii) microtubules. The three cytoskeletal filaments are well differentiated but also physically interconnected [63, 64].

**Microfilaments** Microfilaments are polymeric protein fibers that proliferate the cytoplasm. They are composed of similar actin monomers (homopolymers) and are 8 nm in diameters [65]. In neurons 4-5% of total proteins are actin [66], so that there are about $3 \times 10^5\,\mu$m of filament per cell [65].

In mature axons a network of actin filaments provides stability by forming highly ordered three dimensional networks or closely packed bundles. The bundles provide shape to the cytoskeleton while networks provide gel-like environment. The overall shape of the actin bundles have been reported to depend on the confining space. On relatively free spaces, they form loose isotropic network while in confined geometries the bundles are packed and take the shape of the confining structure. They also act as a substrate for other proteins and facilitate ATP transport.

**Neurofilaments** Neurofilaments are coiled alpha helical structures. Individually they have a diameter of around 10 nm but are commonly found in bundles comprising of about 40 filaments. It is estimated that in a typical mammalian cell neurofilaments have a total length of about $10000\,\mu m$ [68].

Neurofilaments are mostly associated with axonal caliber with increased neurofilament density in thicker axons. The separation between neighboring bundles has been reported to be around 30 to 40 nm [69]. As such they are the structural units of the axons.

**Microtubules** Microtubules are heteropolymers of two globular proteins, tubulin alpha and tubulin beta [70]. To form a microtubule, the units self-assemble and form

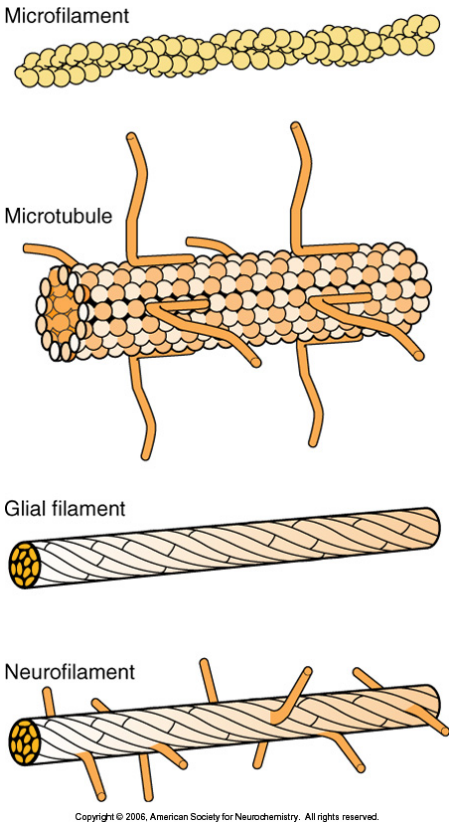



FIGURE 4.2: Cytoskeletal filaments in the brain WM. Each filament structure is long and contains many units of globular proteins. The Microtubules and neurofilaments also have wisp like material that extend laterally. Image reproduced from Siegel et al. [67]

a rigid, hollow tube. A single microtubule has 13 protofilaments in a tubular array. The inner diameter of the cylinder is around 13 nm and the outer diameter is about 25 nm. In a typical culture, Hellier et al. have reported to grow about 150 microtubules with 50 to 100 $\mu$m in length [71]. Even in hippocampal white matter (where fibers are less straight than true WM regions), the microtubules lengths have been reported to be 20 $\mu$m [72]. The outer sides of the microtubules contain hair-like protein projections that could be many nanometers long [69]. Like microfilaments, the microtubular structure is also highly dependent on the cell shape. In typical spherical cell microtubules assemble radially, but in axons they align parallel to the axons.

Microtubules are transport highways of organelles. Within five hours of application, colchincine, a microtubule disassembling drug, has been reported to stop axonal transport [73, 74]. Similarly, introduction of colchincine in olygodendrytes decreases the entry of proteolipid proteins into myelin [75]. In garfish, it has been shown that depolymerization of microtubules with vinblastine reduced the ADC of water in both parallel and perpendicular direction [76]. In rats treated with methylmercury, a decrease in microtubular concentration resulted in an increased water ADC parallel to the axons [77].

The structural integrity of microtubules could also play an important role in diffusion of water inside the tubular lumen. Unlike other cytoskeletal filaments, microtubules are actually hollow cylinders. Considering the inner diameter of 13 nm, and 1 nm think hydration layer, almost a third of the water population inside the microtubules is hydration water. Furthermore, recent studies show that the lumen of the axonal microtubules are filled with particles about 6 nm in size and separated at an average distance of 15 nm [78, 79]. The presence of such particles would also drastically reduce the mobility of water inside the tubular lumen. Water molecules residing inside the microtubules cannot easily escape outside. It is therefore plausible that the microtubular water has relatively slow diffusion in both perpendicular and parallel direction.

## 4.2 Mesoscopic Organization of White Matter

Different composition and relative fractions of protein, lipid and water produce a variety of cellular substructures. However, the basic features are similar for all cellular environments. Membranes are rich in lipid and membrane proteins. They help to differentiate cells and cell organelles from their surrounding and maintain cytoplasmic environment. Membrane-bound cell organelles like mitochondria or lysosomes have their own biochemical milieu. The nuclei are the central units. They store chromosomes that are protein based structures made of DNA or RNA. Other tiny structures like ribosomes, golgi apparatus are mostly protein assemblies. Endoplasmic reticulum is a system of internal membranes found all over the cell. It is relatively rare in axons but can be found in dendritic regions. The cytoplasm itself is composed mostly of water, but the scaffolding of cytoskeletal filaments provides rigidity.

The macromolecules pervade not only the cell cytoplasm and membrane but also the outside in the extracellular matrix. The extracellular matrix is full of polymeric carbohydrates and protein filaments. The presence of myriads different substructures in a small space allows only little separation between neighboring macromolecules. The tight-fitting of numerous small structures in a small cellular volume is commonly referred to as macromolecular crowding.

Macromolecular crowding and its relationship to cellular kinetics have been well studied [80, 81]. Crowding hinders free movement of water molecules present in the cytoplasm. Furthermore, the presence of these macromolecule also result in hydration layers where water molecules are known to have local ordering and decreased diffusivity. The macromolecular crowding is largely responsible for a reduced water ADC in all biological tissue.

For short diffusion times, crowding should result in a time dependent ADC that reflects the overall surface to volume ratio (Eq. 2.32). But, since the macromolecules are essentially small sub-microscopic entities, and diffusion coefficient of free water at body

temperature is around $3\,\mu m^2/ms$, their effect is grossly averaged within a couple of milliseconds. For longer time measurements, the overall effect of small entities can be treated as an effective medium where diffusion coefficient of water is considerably lower than in bulk [12]. A fuller review of literature dealing with macromolecular crowding and reduced diffusion can be found in a paper by Le Bihan [82].

In addition to macromolecular crowding, the CNS is also 'structurally crowded'. The nucleus, cytoplasm, membranes and cell organelles all fit in together in close packed structure. Water molecules residing in some of these gross structures may take longer times to intermingle. The result is a heterogeneous medium with non-Gaussian diffusion propagator where effective medium theory is insufficient to describe water diffusion.

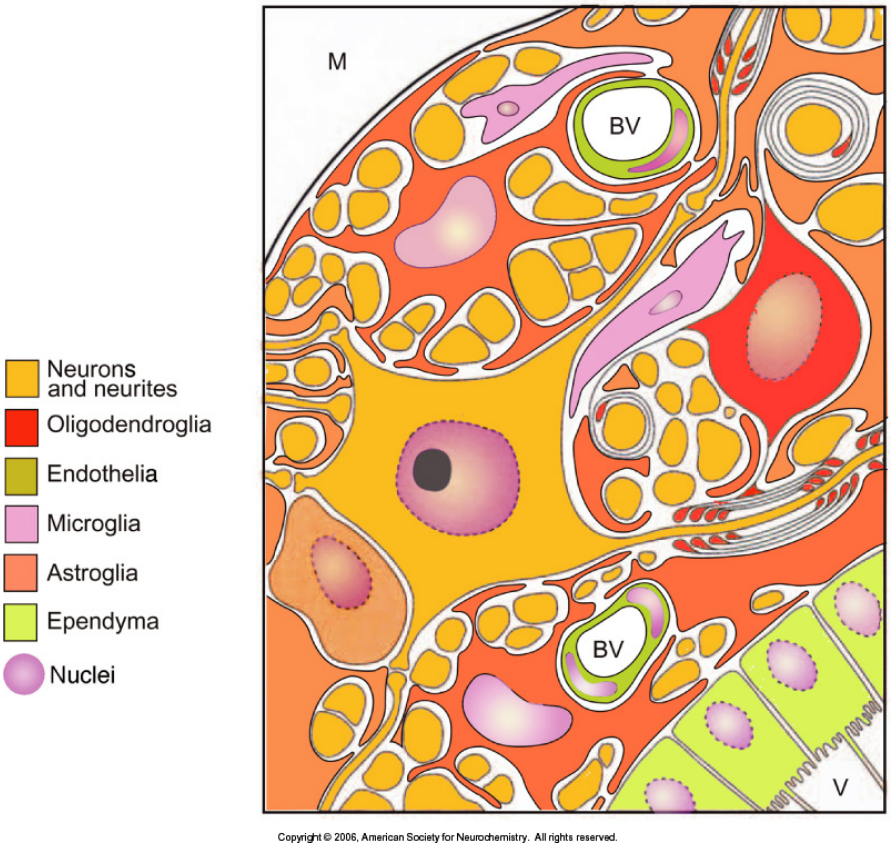



FIGURE 4.3: Structural crowding in the CNS. The major structures are shown with their relationship. The neurons along with axons (shown in transverse sections) are surrounded by astroglia and microglia. The neuronal nucleus is enveloped by membranes and so do other structures. The astrocytes are also connected to blood vessels (labeled BV). The ventricles (V) contain cerebrospinal fluid where diffusion is the most unrestricted. Even the extracellular space ( white spaces) is actually a matrix. Image reproduced from Siegel et al. [67]

WM plays a major role in establishing electrical connectivity; and like all connective tissues (muscles, cartilage) the elongated structures play an important role in its function. In WM, most cellular structures are somewhat found in their elongated form.

The cytoplasm is found in the cylindrical axonal lumen. The mitochondria is naturally cylindrical and resides inside the axon with its principle axis more or less aligned with the axonal axis [83]. The cytoskeletal filaments are also found aligned along the same preferred axis. The myelin membranes are technically part of the oligodendrocytes but they wrap around the axon and are elongated structures. Of course, there are other cellular structures like oligodendrocytes and macroglia which are not necessarily elongated. But compared to the gray matter, anisotropy is the most striking WM feature. The extra-cellular space basically acts as a gap between neighboring axons and still follow the cylindrical symmetry.

This structural anisotropy is also reflected in water diffusion. Water diffusion in WM is largely anisotropic and non-Gaussian. The time taken for water to diffuse from one mesoscopic structure to another could be on the order of the observation time in a diffusion NMR experiment. It is the organization of these mesoscopic structures that is relevant in understanding observed diffusive behavior in the WM.

### 4.2.1 Axon

An axon is prolongation of a neuron that is responsible for electrical signal conduction. The protrusion is filled with cytoplasmic liquid and separated by its own axonal membrane. Axonal lumen contains cytoskeletal filaments, mitochondria, smooth tubular endoplasmic reticulum along with a few lysosomes, and vesicular bodies. Most of these structures are oriented along the longitudinal axis of the axon. While microtubules predominate in unmyelinated axons, as the axons increase in size and myelination, homogeneous arrangement of microtubules are disrupted by neurofilament aggregates.

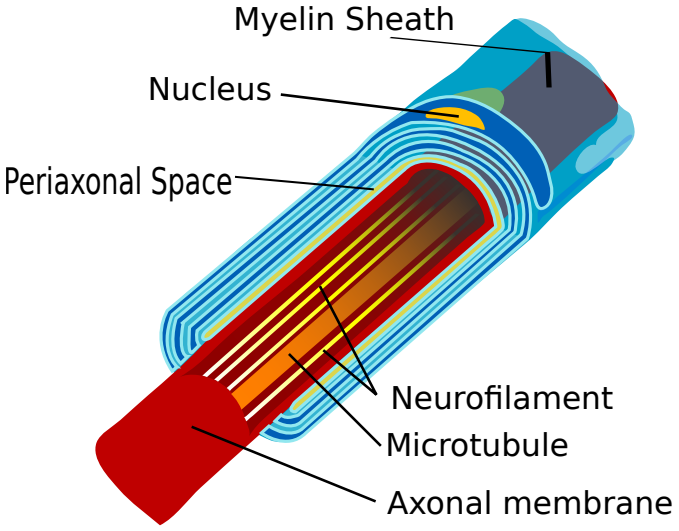



FIGURE 4.4: Schematic of an Axon showing the cytoskeletal filaments in the lumen, axonal membrane, periaxonal space and the myelin layers. The nucleus of the myelin forming cells is also shown.

Axonal water is enclosed within the axonal membrane also known as axolemma which is about 8 nm thick. It is a typical lipid bilayer that has low water permeability. Measurements at short diffusion times could possibly differentiate the axonal water from water in the adjacent periaxonal space (see Fig. 4.4). For longer observation times, such differentiation largely depends on permeability of the axonal membrane that separates axonal water from periaxonal water lying next to it.

Diffusion inside axons have been reported in excised lamphrey spinal cord [84]. Since the ADC was found to be similar in parallel and perpendicular direction, it was argued that only axonal membranes were responsible for the apparent diffusion anisotropy in axons. If axonal membrane plays the major role in diffusion anisotropy then axonal membrane might be quite impremeable to water molecules.

It is likely that the residence time of water associated with tiny structures are so small that their effect is time-averaged and results in a reduced ADC compared to free water. One exception to this is water residing inside the microtubules. As mentioned before, the water inside the microtubules could have a long enough residence time to survive time averaging even at times accessible in a diffusion NMR measurement.

### 4.2.2 Periaxonal Space

The periaxonal space is 10-20 nm wide, extracellular space trapped between the myelin sheath and the axonal membrane [85]. It has been reported that diffusion of tracer particles between outer regions and the periaxonal space is possible but has to follow a tortuous path and is in the order of hours [86]. This space is filled with myelin associated glycoprotein (MAG) which is associated with periaxonal membrane of the inner myelin layer.

Water diffusion inside periaxonal space should be limited and anisotropic. While water can move up and down along the axonal axis, its radial displacement is limited around the axonal membrane. Even in the axial direction, diffusion is limited: water molecules are squeezed between the axonal membrane and the myelin membrane both of which contain lipid layers. Due to lipid membranes on both sides, water molecules in periaxonal space have little chance to intermingle with the outside. Some exchange between periaxonal water and axonal water (except in the microtubules) might be expected. However, it would be even more difficult to mix with myelin water that lies on the other side of the periaxonal membrane.

### 4.2.3 Myelin

Myelin is an extension of glial plasma membrane that radially wounds around the periaxonal space. It is is composed mostly of lipids, protolipid protein (PLP) and myelin basic protein (MBP). The layers of myelin are well known for their insulating properties

that allow up to a 100 fold increase in conduction velocity. Most of the myelin in WM is compact myelin.

As the oligodendrocyte processes wind around the axon, the compact myelin layers are formed. As oligodendritic processes wrap around the axon, the membranes in the compact myelin are stacked. Therefore, a layer is actually two adjacent lipid bilayers of membrane with a 2.5 nm wide cytoplasmic layer in between. Between two adjacent windings there is another 4 nm wide extracellular space. The membranes and the cytoplasm contain MBP that glues the two layers together. The PLP is abundant in extracellular space and keeps neighboring layers intact. The repeat distance of myelin layer in brain WM is around 16.6 nm [87].

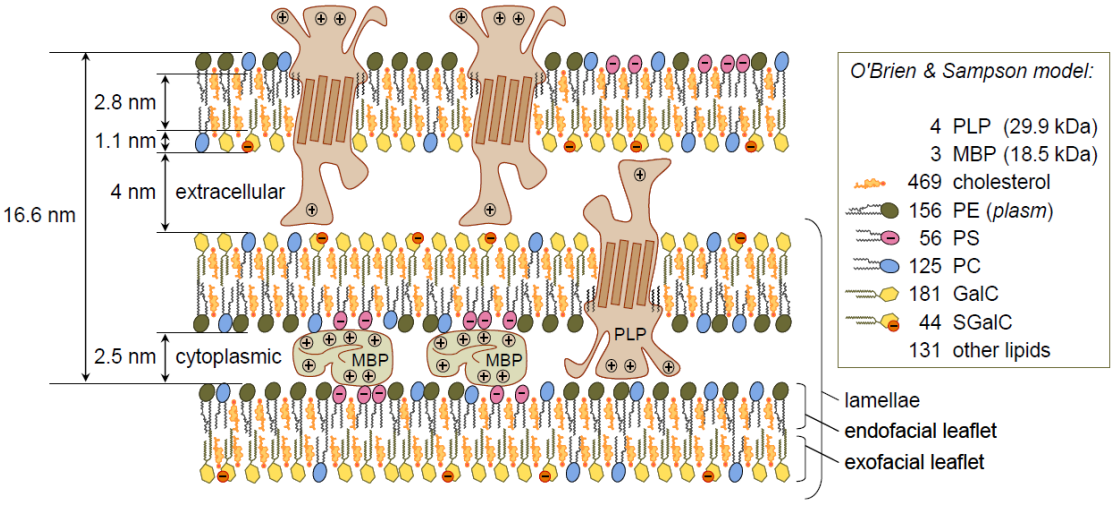



FIGURE 4.5: Model of compact myelin based on experimental electron profile density derived from small-angle X-ray scattering data of human brain surgical biopsies [88]. The myelin composition is adapted from O'Brien and Sampson [49] and recent proteomics study of myelin [89]: 4 proteolipid proteins (PLP), 3 myelin basic proteins (MBP), 125 phosphatidylcholines (PC), 181 galactosylcerebrosides (GalC), 44 sulfatides (SGalC), and 131 other lipid. Cartoon adapted with permission from Dr. Christian Labadie.

The myelin lipid bilayers contain high fractions of cholesterol and sphingomyelin. About 30% of lipid content in myelin is cholesterol [90]. Such high fraction of cholesterol dramatically reduces the permeability of the myelin membrane [50].

For water molecules residing between the myelin layers, moving through multiple layers of bilipid layer would be an almost impossible feat, but they can still move parallel to the axons and radially around the axonal lumen. However, almost all water in the 2.5 nm cytoplasmic space and more than half the water in the 4 nm extracellular space is hydration layer. This would drastically reduce diffusion properties of myelin water. The presence of MBP and PLP also reduce water mobility. At long time it is expected that myelin water ADC is rather small in both parallel and perpendicular direction.

### 4.2.4 Node of Ranvier

At periodic intervals the axolemma is relatively uncovered in sites known as nodes of Ranvier. These are sites of chemical exchange responsible for boosting the action potential. The nodes are 1-2 $\mu m$ long. Figure 4.6 shows an electron micrograph of the node of Ranvier.

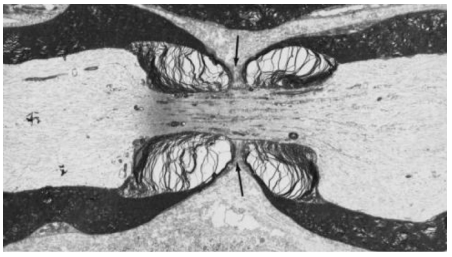



FIGURE 4.6: An electron micrograph of the node of Ranvier. The axon diameter is decreased and axolemma comes in contact with astroglial space. Picture taken from Siegel et al. [67]

The nodes of Ranvier are proliferated with voltage gated Na+ channels [91]. These channels readily allow the passage of protons. Only microtubules are continuous across the node of Ranvier. Otherwise, the nodes are easy site of water exchange between axons and the astroglial space. However, Ranvier nodes are relatively rare with a periodicity of at least 50 $\mu$m and usually more than 150 $\mu$m [92]. Therefore, the axonal lumen remains relatively isolated during the observation time possible in NMR.

### 4.2.5 Glial cells

There are three kinds of glial cells in the WM. They are astrocytes, oligodendrocytes and microglia.

The astrocytes are the biggest of the three with somatic diameters around 10-12 $\mu m$. In WM, they are more or less randomly oriented, oblongated cell bodies[93]. The processes are smooth, thin, have relatively less branching, and longer in the opposite directions along the oblongation of the cell body and extend more than 100 $\mu m$ [94] These processes are unbranched in WM and are densely filled with neurofilaments and few microtubules. While most processes have their endfeet in contact with axons at nodes of Ranvier, one process has its endfoot in contact with the blood capillaries [95]. The water channels, aquaporins, are expressed in the endfeet. Astrocytes are connected to other astrocytes and oligodendrocytes through gap junctions.

In WM, oligodendrocyte somas are arranged in rows along the axons and are called interfascicular oligodendrocyte. The cell bodies are 6-8 $\mu m$ in diameter and their processes wrap around the axons to form myelin layers. The oligodendritic processes are

extended more or less isotropically. Microtubules are present throughout the oligodendritic cell bodies as well as in large and small processes. They are randomly oriented in the cell body but are aligned parallel to the processes [96].

Microglia clean up the tissue environment by engulfing dead cells and waste. The microglia in WM is mostly found in ameboid form. They are allowed to move in the WM and look for any damage in its architecture. Microglia are therefore smallest in size.

The lipid permeability of cell membranes of both astrocyte and oligodendrocyte have a low membrane permeability, but these membranes lack cholesterol and sphingomyelin that are found in the myelin. Furthermore, unlike myelin, they are also just a single bilipid layer. However, it is likely that these membranes are as impermeable as axonal membranes. Considering that axonal membranes have been implicated as the primary source of anisotropy in brain WM, it is plausible that water inside the glial processes also have anisotropic properties even for long diffusion times.

For short time diffusion measurements, the processes can be approximated by cylinder distributed isotropically in all directions. Approximating diffusion as a tensor for each cylinder, the average diffusion propagator would be the integral of a single tensor over all of the volume [97]. The diffusion decay curve would be isotropic but non-monoexponential. The normalized signal can be written as [97]

$$S = e^{-bD_\perp} \int_0^1 e^{-bD'_\parallel x^2} dx = e^{-bD_\perp} \frac{\sqrt{\pi}}{2\sqrt{bD'_\parallel}} \operatorname{erf}\left(\sqrt{bD'_\parallel}\right); \tag{4.1}$$

where $D_\perp$ and $D_\parallel$ are the ADC perpendicular and parallel to the processes, $D'_\parallel = D_\parallel - D_\perp$ and $erf$ is error function. When $bD'_\parallel << 1$ the ADC is equal to $(D_\parallel/3 + D_\perp)$. For higher b-values the signal shows non-monoexponential behavior. At $bD'_\parallel >> 1$, the error function approaches 1 and the signal close to some b-value $b_0$ can be approximated as

$$ln(S) = -b_0 D_\perp + ln\left(\frac{1}{2}\sqrt{\frac{\pi}{bD'_\parallel}}\right) - \frac{1}{2}\left(\frac{b}{b_0} - 1\right) \tag{4.2}$$

Since the slope of the curve above is dependent on the b-value. If we loosely define ADC as the slope of the log of the signal at each b-value, it is dependent of the b-value of choice such that ADC(b)= $D_\perp + 1/2b$.

But, measurements at long times may be different. Most processes have very thin diameters [98] such that for most diffusion MR measurements $D_\perp$ is perhaps a very small number close to zero. In such a case, directly from Eq. 4.1, the signal at high b-factors is a function of $1/\sqrt{b}$

$$S(b|\sqrt{bD_\parallel} >> 1) = \frac{\sqrt{\pi}}{2} \frac{1}{\sqrt{D_\parallel}} \frac{1}{\sqrt{b}} \tag{4.3}$$

In practice the error function quickly approaches 1, and the equation is valid when $bD_{\parallel} \geq 6$. In the absence of any other components the slope of the fit would reveal $D_{\parallel}$. However, care should be taken that at long times when molecules move a long distance and could exchange with water from other substructures. This could have some effect on their volume fractions.

Similar models of three dimensionally distributed cylindrical processes has been used to model dendritic diffusion in GM and diffusion in respiratory airways [99, 100].

### 4.2.6 Vasculature

Human brain consumes around 20% of oxygen requiring a blood supply of about a liter per minute. The flow of blood moving through the vessels could be mistaken for fast diffusion. In most case the 'pseudo-diffusion' due to vascular contribution is more than twice the ADC in free water and its contribution can be readily separated.

## 4.3 Diffusion Models in WM

Since the early days of diffusion NMR, there has been a keen interest in understanding molecular mobility in biological samples. NMR stands out among all other techniques because it is non-invasive and probes samples with minimal interference. It can be used to study biology as it happens.

The displacement of spin carrying water molecules is in the order of micrometers. Diffusion MR measurements can allow us to sense these micro-structures without having to cut the sample. This sensitivity towards microscopic geometry has established diffusion-weight magnetic resonance (DW-MR) as an invaluable clinical and research tool. However, the relation between the diffusion weighted signal and the underlying WM tissue micro-structures has so far only been qualitatively understood.

Since the inception of this field, understanding diffusion MR signal in biological tissue has been known to be a complex inverse problem. On one hand, diffusion propagator from two different systems might be very similar; propagators for most heterogeneous systems have a Gaussian like central distribution with different tails. A major part of the propagator therefore just reveals the ADC. On the other hand, exponential nature of signal decay allows for a very small window to distinguish between signals that come from very different origins.

This thesis attempts to understand the physical basis of non-monoexponential diffusion decay curves in brain WM.

The following section is a brief description of the review of relevant studies. Three models are explained that are representative of the major directions that the field has taken over the years.

A natural first approach to this problem is to describe the signal. For simple fluids, Eq. 3.54 conveniently described the signal behavior as an exponential decay related to the b-value and the diffusivity. At low b-values (bD$<$ 1) diffusion decay curve still resemble an exponential decay. Assuming Gaussian decay, the curves can be used to estimate the ADC is related to the variance of the spins. If the diffusion is Gaussian-like in different directions but anisotropic, then it is possible to describe the signal with the help of a diffusion tensor. Diffusion tensor imaging (DTI) is the most widespread model used to interpret data from diffusion weighted magnetic resonance (DW-MR) technique.

### 4.3.1 Diffusion Tensor

Diffusion tensor was previously defined with Eq. 2.8. It can account for anisotropic diffusion but still assumes a homogeneous Gaussian diffusion. Diffusion tensor as a concept predates the discovery of NMR, but it has been systematically used in diffusion-MR studies only after the diffusion tensor model proposed by Basser et al. [101].

The diffusion coefficient in any arbitrary direction $\vec{\mathbf{r}} = [x, y, z]$ can be expressed as

$$D(\vec{\mathbf{r}}) = \begin{bmatrix} x & y & z \end{bmatrix} \begin{pmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{yx} & D_{yy} & D_{yz} \\ D_{zx} & D_{zy} & D_{zz} \end{pmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \tag{4.4}$$

Since the diffusion tensor is symmetric [7], the equation above can be expanded as

$$D(\vec{\mathbf{r}}) = x^2 D_{xx} + y^2 D_{yy} + z^2 D_{zz} + 2xyD_{xy} + 2yzD_{yz} + 2zxD_{zx} \tag{4.5}$$

The linear equation above can be solved to calculate the diffusion tensor if the diffusion coefficient in six independent directions are known. The tensor can be reduced to 6 independent elements (3 diagonal and 3 off-diagonal). In practice calculating the diffusion coefficient also requires evaluating signal without any diffusion weighting. Thus, DTI requires at least 7 measurements.

In the previous chapter, DW-MR signal for free diffusion was described as a exponential decay in terms of b-value. This b-value is proportional to the square of the gradient amplitude. DTI uses this property to estimate the tensor. Following Eq. 3.54 and Eq. 4.5, one can directly estimate the diffusion tensor from the logarithm of the signal obtained.

$$ln(S) = ln(S_0) + \gamma^2 \delta^2 (\Delta - \delta/3) \vec{\mathbf{g}} \mathbf{D} \vec{\mathbf{g}}' \tag{4.6}$$

where $\mathbf{D}$ is the diffusion tensor, $S$ is the spin-echo signal for each measurement and $S_0$ is the spin echo signal without any diffusion weighting. In practice the tensor is estimated as a oversampled linear problem, where the number of diffusion directions is far more than the required minimum of 7. The signal follows a Rician noise distribution therefore,

the linear inversion is only to be used for high SNR data. The diagonalization of the diffusion tensor gives its eigensystem. Since the tensor is positive definite we expect all the eigenvalues to be positive.

The diffusion tensor formalism has been widely used to report on WM property and integrity. Scalar measures derived from the tensor allow direct comparison between tensor systems without depending on its directionality. Scalar indices also help visualize different aspects of the tensor in a single map. The basic scalar measures are the eigenvalues themselves and the trace of the tensor. The trace is equal to the sum of the three eigenvalues and the three dimensional ADC or the mean diffusivity is one third of the trace.

An important rotationally invariant index is the Fractional Anisotropy (FA). FA reflects the variance in the tensor eigenvalues and is a measurement of the anisotropy of the system. It is defined as.

$$FA = \sqrt{\frac{3}{2}} \frac{\sqrt{(\lambda_1 - \hat{\lambda})^2 + (\lambda_2 - \hat{\lambda})^2 + (\lambda_3 - \hat{\lambda})^2}}{\sqrt{\lambda_1^2 + \lambda_2^2 + \lambda_3^2}} \tag{4.7}$$

where the $\lambda$s are the three eigenvalues of the tensor and $\hat{\lambda}$ is the ADC. The isoprobability surface (see Fig. 2.1) of the diffusion tensor has long been used to visualize the anisotropy of brain white matter. The FA always scales between 0 and 1 and has been found to be relatively insensitive to noise [102]. It is the most widely reported anisotropic metric.

Similar to FA different ratios between the eigenvalues have been used to describe the behavior of the tensor. In this thesis, we have extensively used coefficients of linearity $\left(c_l = \frac{\lambda_1 - \lambda_2}{\lambda_1}\right)$, planarity $\left(c_p = \frac{\lambda_2 - \lambda_3}{\lambda_1}\right)$ and sphericity $\left(c_s = \frac{\lambda_3}{\lambda_1}\right)$ where $\lambda_1$,$\lambda_2$, and $\lambda_3$ are the first, second, and third eigenvalues of the diffusion tensor. These coefficients describe proximity of the tensor to a line, plane and sphere [103].

The most straightforward interpretation of DTI model would probably be at very short diffusion times. At very short times, when diffusion length is very small relative to the enclosed space, most of the diffusing particles do not feel the boundaries. In this short time regime, we expect the measured time dependent ADC to be inversely proportional to square root of time and related to the surface to volume ratio [11, 104]. Evolution of the tensor from isotropic to the anisotropic regime would be extremely helpful in understanding the physical basis of such behavior. However, experiments in WM have shown a non-monoexponential signal decay even for the shortest diffusion times. Presence of non-monoexponential decay limits the applicability of DTI model.

Another way to use the DTI model would be at very long times. WM is microscopically heterogeneous - the shape and size of different tissue substructures act as aggregate boundaries. The boundaries are most relevant when the average diffusion length ($\sqrt{tD(t)}$) is close to the size of the boundaries. [12, 105]. With increased diffusion time,

as more and more particles encounter the boundaries, the shape and the width of the propagator is sensitive to the structure and density of the tissue micro-structures. At very long times, the diffusing particles probe all of the heterogeneous spaces and the resulting diffusion propagator loses information about the microscopic ultra-structure and only retains the information of the gross tissue shape (macrostructure). Due to central limit theorem, the propagator once again will approach a Gaussian and DTI model would resume its applicability.

The diffusion tensor deals exclusively with the ADC in each direction and implicitly assumes Gaussian diffusion. Its strength lies in its simplicity and close relationship to tissue parameters. The DTI model assumes that anisotropy seen in WM is more a reflection of its ordered tissue structure rather than the molecular hydrodynamics. The eigenstructure of the tensor can be used to infer geometry of the diffusion compartments. For WM with single coherent fiber bundles, the primary eigenvector (with largest eigenvalue) is parallel to the main fiber orientation [101]. Similarly for dispersed fibers, the minor eigenvector is perpendicular to the plane of fiber dispersion [106].

The study of WM anisotropy however predates the widespread use of DTI in DW-MR measurements.

Beaulieu and Allen performed measurements in non-myelinated garfish axons and reported reduced anisotropy [76]. This lead them to conclude that diffusion anisotropy is related to axonal orientation but myelination nevertheless played a major role in enhancing it. In their other study in giant squid axons, they showed very small anisotropy in the intraaxanol water [107]. They also reported that other intra-axonal microscopic features like neurofilaments and the macrotabecular lattice were not primary contributors to the observed anisotropy in WM.

A significant contribution to understanding WM anisotropy was a study by Takahashi et al. [84], who performed micro-imaging in excised lamphrey axons. With an in-plane resolution of $19 \times 19 \mu m^2$ they were able to study diffusion of intra-axonal water. At a diffusion time of (11 ms ) Takahashi et al. reported that no diffusion anisotropy was observed in the biggest axons ($\approx 40 \mu m$) but smaller axons showed progressively greater anisotropy. This study established that the WM anisotropy was closely related to its orientation. Measurements on myelin deficient rats [108], rat pups [109] and also on human infants [110] have reiterated such claims. In *ex vivo* samples, measurements by Tuch et al. showed that conduction and diffusion in WM are closely related [111].

The most noticeable structure in WM is the myelinated sheath. Since the early days of DW-MRI, diffusion in brain white matter has been closely linked with the presence of myelin. It is reasonable to expect diffusion measurement in brain WM to be influenced by axonal diameter and permeability of the myelin layers. Le Bihan et al. measured the diffusion time dependence for WM in human brain *in vivo* and reported no significant change in ADC for diffusion times ($\Delta$) between 16 and 79 ms [112] . Time invariance of

the ADC led them to conclude that the ADC attains its asymptotic long time limit within 16 ms. Following this ansatz, they concluded that myelin is quite permeable to water molecules and calculated the permeability coefficient to be $1.6 \times 10^2$ cms$^{-1}$. Moonen et al. also reported time invariant ADC for measurements in cat brain measuring for diffusion times from 50 to 2000 ms [113]. This was also consistent with another study by Pfeuffer et al. on rat brain where they suggested that the intracellular diffusion exchange time was around 15 ms [114].

While the study of diffusion anisotropy was underway, several findings of non-monoexponential diffusion decay were noticed. In garfish, Beaulieu and Allen noticed a clear deviation from a monoexponential decay at high b-factors [76] that could be fitted with three exponentials. Similar triexponential fits were also reported by other studies employing very high b-factors (>50 $ms/\mu$m$^2$) [115, 116] Using single- and double-quantum-filtered diffusion spectroscopy Assaf and Cohen also reported distinct water populations in rat brain [117]. Several other studies have shown that diffusional decay curves in brain WM significantly deviate from a single exponential decay [118–122].

### 4.3.2 Gaussian Mixtures

An easy way to fit non-monoexponential decay is to assume multiple diffusion pools, where each diffusion pool is still a Gaussian but the pools do not mix with each other (Section,3.6.1). In a Gaussian mixture model the signal decay is described as

$$S(b, \vec{D}, \vec{F}) = S_0 \sum_{j=1}^{2} \vec{F}_j \; e^{-b\vec{D}_j} + \varepsilon_i \tag{4.8}$$

Here $\vec{D}$ is now a vector including both diffusion coefficients and $\vec{F}$ is the vector with volume fractions for each pool [123]. When diffusion decay is measured in a single direction, the simplest Gaussian mixtures is a biexponential. The bi-exponential model is a simple extension of the mono-exponential model to two pools of water present. The tri-exponential model further assumes three pools.

### 4.3.3 Stanisz Model

The first analytical modeling of non-monoexponential diffusion decay curves in WM was performed by Stanisz et al. [124]. The basic framework for their description was obtained from electron microscopy of the WM specimen. Ruminating on diffusion anisotropy and non-exponential diffusion decay, they modeled axons as long cigar shape ellipsoid and other extra-axonal but cellular structures (including all of oligodendrocytes, microglia and astrocytes) as spheres. Diffusion in extracellular water was also modeled to be anisotropic. This three compartment model comprehensively describes diffusion decay

curves as arising from cumulative effect of cellular dimensions and membrane permeabilities. Using Kärger's exchange equation (See Section 3.6.2.2), permeabilities of cellular membranes as well as myelin layers were also taken into account. Both these compartments were considered to be in exchange with the extracellular space. A summarized description of Stanisz's model can be seen in Fig. 4.7

Stanisz et. al, MODEL

Axons as prolate ellipsoids



Glial cells as spheres



$a_t$ = length of axonal major axis
$a_t(\perp)$ = radial length of axons
$V_t$ = Volume fractions of axonal water
$P_t$ = permeability of axonal myelin

$a_s$ = diameter of glial cells
$V_s$ = Volume fraction of glial water
$P_s$ = permeability of glial membranes

$D_e$ = Extracellular diffusion coefficient
$D_i$ = Intracellular diffusion coefficient

DEPENDENT PARAMETERS

**Apparent Diffusion Coefficients**:

$$D_e = D_e(1 - V_s - V_t)^M$$

where, $M = \frac{0.5V_s - 0.5V_sL + V_tL}{(1-L)}$

$$L(\|) = ln\left(\frac{2a_t(\|)}{a_t(\perp)}\right)\left[\left(\frac{a_t(\perp)}{a_t(\|)}\right)^2\right]$$

$$L(\perp) = 0.5 - 0.5L(\|)$$

$$D_i = -\frac{1}{b}ln\{2\frac{1-cos(ql)}{q^2l^2} + 4q^2l^2\Sigma exp\left(\frac{n^2\pi^2D_i}{l^2}\right)\frac{1-(-1)^n cos(ql)}{(q^2l^2-n^2\pi^2)^2}\}$$

here, $q = (\gamma\delta g)$
for glial cells: $l = a_s/2$
for axons: $l(\|) = a(\|),\ l(\perp) = a(\perp)/2$

**Exchange from Permeabilities**

$k_s$ = glial water to extracellular space
$= \frac{6P_s}{a_s}$

$k_t$ = axonal water to extracellular space
$= \frac{3P_t}{a_t(\perp)} + \frac{1}{\varepsilon}arcsin(\varepsilon))$

$\varepsilon$ = eccentricity of ellipse formed by cross section of ellipsoid through symmety axis

$$\frac{dM_t}{dt} = -q^2D_tM_t - k_tM_t + k_eM_e$$
$$\frac{dM_s}{dt} = -q^2D_sM_s - k_sM_s + k_eM_e$$
$$\frac{dM_e}{dt} = -q^2D_eM_e - k_eM_e + k_tM_t + k_sM_s$$

FIGURE 4.7: Stanisz Model

Some of the parameters obtained by their model were found to be close to their actual values. Most interestingly, axonal diameter estimated to be around 2.5 $\mu$m was similar to values obtained from electron microscopy. Similarly diameter of the glial cells was also found be around 3 $\mu$m. Considering that glial cells also include oligodendritic and astrocytic processes, the effective size of these structures could have exceeded their somatic size. Exchange time of the intra-axonal water to the extra-cellular water and intra-glial water to extra-cellular water was estimated to be around 62 ms and 30 ms

respectively (exchange rate of 16 $s^{-1}$ and 33 $s^{-1}$). However, a 40%, volume fractions that was assigned to the extracellular compartment was quite unphysical.

Although simplistic compared to reality, the Stanisz's model still requires quite a few parameters. The necessity of such an elaborate model might reflect inherent complexity of tissue structure, but it could also be due to a mismatch; ignoring features that are relevant and modeling features that are not. The model requires an estimate of axonal length which helps to account for an upward moving decay curve with increasing $\Delta$, but it does not have a clear physical meaning. Similarly, including exchange of both axonal and glial water with small extracellular space made the equations overloaded with fitting parameters.

The difficulty in applying Stanisz model is apparent for measurements at longer diffusion times. Similar to Stanisz et al., Assaf and Cohen [122] measured excised bovine optic nerve, but while Stanisz et al., employed diffusion times ($\Delta$) between 8 and 30 ms, measurements by Assaf and Cohen were performed for $\Delta$ up to 350 ms [122]. At such long observation times, Stanisz model would expect a close to monoexponential signal decay. However, Assaf and Cohen reported a clear non-monoexponential decay curves for both parallel and perpendicular directions [122].

Though Stanisz et al., described diffusion decay with three compartments, subsequent measurements were mostly focused on biexponential nature of the curves [119, 120, 122]. Measurements were performed in single direction where a two compartment biexponential still produced a good fit to the data. A two compartment system was easier to model and to relate to intra and extra-axonal compartments. Mulkern et al., measured diffusion decay curves in human brain in vivo and observed biexponential diffusion behavior in both GM and WM [120]. However, due to limited gradient strength the measurements did not deal with direction dependence. Clark and Le Bihan also observed a biexponential decay in human brain WM *in vivo* and could also measure distinct decay curves along and perpendicular to the fiber orientation [119]. Their biexponential analysis yielded different volume fractions for fast and slow diffusion compartments depending on the measurement direction. Volume fraction of the slow component was found to be around 14% parallel to the fibers but was about 40% in the perpendicular direction. The volume fractions of the two components reported by Assaf and Cohen were similar to those reported by Clark and Le Bihan [122].

Human *in vivo* measurements suggested that the two components had similar $T_1$, $T_2$ and magnetization transfer ($MT$) properties [125, 126]. The early measurements were helpful in revealing non-monoexponential decay curves in WM, however, resolving apparent discrepancies in relative volume fractions and exchange times between compartments were still open.

At this time, the research focus slightly shifted towards using this non-monoexponential behavior as an MR contrast tool to delineate tissue structures and pathologies. Inglis et

al., used a biexponential diffusion tensor in rat spinal cord and claimed proper visualization of extra- and intracellular water compartments [127]. The maps of fast and slow components showed clear distinction between WM and GM. Similar bi-tensor model was also applied in human brain in vivo [128, 129]. These measurements showed that for a single tensor analysis mean diffusivity is similar between GM and WM but the ADC of the slow component showed clear difference between the two. The ADC of the slow component was found to be much lower in WM than GM. Though the bi-tensor model resolved wrong volume fractions, these measurements were thought to be useful for medical contrast and tissue characterization. Using high b-value DWI to categorize neuropathology is still an ongoing research issue [130–132].

### 4.3.4 CHARMED Model

In 2004, Assaf et al. [133] proposed a model of diffusion in WM where the two diffusion compartments were intra-axonal and extra-axonal. In this picture, myelin wall is considered to be almost impermeable such that water inside and outside the axons are clearly compartmentalized. All extra-axonal spaces including glial cells and extracellular spaces were combined to a single compartment.

The physical interpretation of two distinct diffusion components is diametrically opposite to the previously considered relatively fast exchange argument. In this new scenario, myelin layers are considered highly impermeable such that diffusion inside the axon is highly restricted radially but relatively free along the axonal axis. It is consistent with the expected reduction in myelin permeability due to high cholesterol and sphingomyelin content [50]. In the perpendicular direction, the maximum displacement for intra-axonal water is governed by the inner diameter of the axons. Radially, intra-axonal water could only move a few microns before hitting the impermeable myelin and reaching its maximum displacement. Once this limit is reached, displacement for the intra-axonal water cannot further increase over time. Radial intra-axonal ADC ($D_{IA\perp}$) would be very small and inversely proportional to the diffusion time [134].

The water fraction residing in the extra-axonal space, including glial cells and their processes, is considered to be hindered but not restricted so that after a certain time they reach their steady state extra-axonal diffusion coefficient $D_{EA}$. Since the measured ADC is an ensemble average, $ADC = V_{EA} \times D_{EA} + V_{IA} \times D_{IA}$, where $V_{IA}$ and $V_{EA}$ are the volume fractions of intra and extra-axonal space respectively. When $D_{IA} << D_{EA}$, measured ADC is very close to $D_{EA} \times V_{EA}$.

This model named composite hindered and restricted (CHARMED) [135], resolved two physically meaningful components and circumvented the problem of resolving exchange times. For a homogeneous WM bundle, CHARMED requires a single hindered water pool (extra-axonal) and a uni-directionally restricted water pool (intra-axonal).

Avoiding exchange, parameters can be estimated with a single diffusion time measurement.

The beauty of CHARMED model lies in its exploitation of the cylindrical symmetry. The intra-axonal compartment naturally has cylindrical restrictions and it is reasonable to expect that extra-axonal water diffusion also follows cylindrical symmetry. Due to the underlying cylindricity, the displacement PDFs along the principle axis and the radial direction are independent. Similar to DTI measurements, linear independence between the parallel and perpendicular allows a diffusion weighted signal in any arbitrary direction to be estimated as function of the diffusion weighting in the parallel and perpendicular direction.

A brief pictorial summary of the CHARMED model can be found in Fig. 4.8.

Restricted diffusion in intra-axonal space has been reported for NAA diffusion [136], and for water diffusion in sciatic nerves [122], but it has not been explicitly observed in brain WM. Many attempts to quantify restrictions in brain WM have led to ambiguous results [137, 138]. Assaf and Cohen found evidence of restricted diffusion in bovine optic nerve and rat spinal cord but could not confirm so for rat brain WM [122]. It should be noted that axonal size in nerves and spinal cord is quite different from those in brain WM. A change in diameter could result in big changes in residence time of intra-axonal water. Clark and Le Bihan could not find any evidence of restrictions in human WM *in vivo* [119]. And in formalin fixed porcine nerves, the slow diffusion component showed non-Gaussian behavior only when $\Delta$ was less than 6 ms. [139]

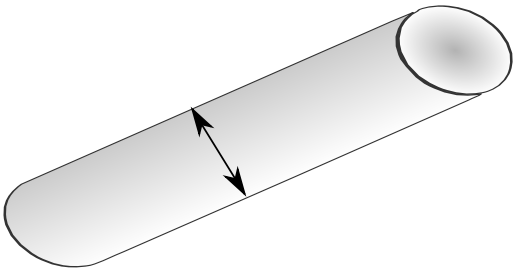
The CHARMED model cleverly treats restricted intra-axonal diffusion as Gaussian. In this sense it is simply a Gaussian mixture model with added constraint that the two tensors have same principle axes (eigenvectors) but with different eigenvalues.

After CHARMED, many studies of non-Gaussian diffusion in WM have primarily focused on diffusion properties perpendicular to the fiber directions [140–145]. These studies implicitly treat layers of myelin as ultimate restrictive barriers and did not consider possible multi-compartmental nature of diffusion data that may be intrinsic to the WM. An extended version of CHARMED model that includes axonal diameter distribution is termed AXCALIBER [146, 147]. In order to estimate axonal diameter AXCALIBER also requires diffusion measurements at different diffusion times.

The CHARMED model can effectively describe decay curves obtained perpendicular to the fiber direction, it falls short in explaining non-monoexponential decays in the parallel direction. In the first paper on CHARMED, Assaf et al., reported noise contribution to be around 10% [133]. This is close to the contribution of the slow diffusion component of 14% that was reported by Assaf et al. [122].

CHARMED MODEL

Axons as long hollow cylinders



Diffusion-smoothed extra-axonal microstructures
Anisotropic Gaussian Compartments

$V_{EA}$ = Volume fraction of extra-axonal water

$D_{EA}$ = Extra-axonal diffusion coefficient

Time independent Diffusion Coefficients

$V_{IA}$ = Volume fractions of axonal water

$D_{IA}$ = Intraaxonal diffusion coefficient

$\parallel$ : Time independent Gaussian Diffusion

$\perp$ : Time dependent diffusion due to restriction

CYLINDRICAL SYMMETRY

Due to cylindrical symmetry for both compartments,
displacement pdfs along the major and minor axis are independent

$$P(\vec{R}) = P_{\perp}(\vec{R}_{\perp})P_{\parallel}(\vec{R}_{\parallel})$$

Diffusion weighted signal (E) at any angle to the major axis can be
expressed as a product of diffusion weighted signal parallel and perpendicular to it.

$$E(q,\Delta) = E_{\parallel}(q_{\parallel},\Delta)E_{\perp}(q_{\perp},\Delta)$$

Eor Extra-Axonal diffusion, $E(q,\Delta) = exp\left(-4\pi^2 q^2(\Delta - \delta/3)q^T Dq\right)$

For Axons,

$$E_{\parallel}(q_{\parallel},\Delta) = exp(-4\pi^2 q^2(\Delta - \delta/3)D_{\parallel})$$

$$E_{\perp}(q_{\perp},\Delta) = exp\left(\left[-4\pi^2 R^4 q_{\perp}^2 D_{\perp}\tau\right](2 - \tfrac{99r^2}{112D_{\perp}\tau})\right)$$

FIGURE 4.8: CHARMED

### 4.3.5 Other Models

In more recent years there has been a resurgence in finding a good fit to the diffusion decay curves. The challenge of effective and realistic modeling is underlined by a recent article [148] comparing 47 different models found in the literature. Most models considered are Gaussian mixture models where different diffusion compartments are individually treated as tensors. The difference between the models only lie in the constraints that are added. For example, a spherical component will be forced to have equal eigenvalues. A cylindrical compartment is required to have the two smaller eigenvalues

to be the same. In the spirit of CHARMED modeling, the diffusion components can also be forced to have the same eigenstructure. Two tensors can also be forced to have one or more same eigenvalues.

### 4.3.6 Higher order models

Similar to the tensor, the kurtosis of the diffusion propagator can also be expressed as a tensor. The kurtosis tensor is a fourth order tensor, with only partial symmetries. It takes 21 parameters to describe such a tensor. Kurtosis tensors are recently gaining more use in DW-MR research. The detailed description of the kurtosis and higher order tensors can be found in Refs. [149] & [150]. The relationship between higher order parameters and the tissue structure is yet to be fully studied. The stretched exponential model also can be extended to allow a tensor of the stretched exponent.

Similar to what DTI is to ADC in one direction, any model of the signal in one direction can in principle be extended to its three dimensional version.

## 4.4 A slight detour

Here, I take the liberty to detour and divide the problem at hand from a philosophical standpoint. A scientific understanding of such complex systems requires the answer for three question.

- How to describe the signal?
- How universal is this description?
- What is the underlying basis of the signal behavior?

The first question is best answered mathematically. Given the current development in mathematics, any continuous curve can be explained with different mathematical forms. Many readily available fitting programs can fit any curve with the desired precision. Increasing the number of fit parameters can give better fit to the data. But increasing parameters makes them prone to fitting noise in the data, making them more unstable. A good mathematical model is the most parsimonious one among many that fit the data well.

The second question deals with how small changes in the measurement affects the description. A good model should be able to describe how small changes affect the data behavior. The change could be made both to the sample or to the measurement. When the mathematical model used to answer the first question has a physical background, the relationship between mathematical and physical parameters is relatively straightforward and the limits of the model could be analytically expected. But, the relationship gets more difficult to decrypt as the mathematical model gets further from physical reasoning.

The third question is more of a why question that tries to understand rather than explain[2]. A physical model would be very helpful to provide answers to how things work the way they do. Fortunately! scientific questions reemerge when the experimental finding surpasses previous understanding.

## 4.5 What is to be done?

Many measurements on WM over the last 3 decades point to three simple things.

1. Diffusion decay curves in WM are non-monoexponential.
2. The curves depend on the angle between the gradient and the fiber direction.
3. The decay curves are relatively stable for different diffusion times.

Given the multitude of different ways data have been previously analyzed, there does not seem to be a dearth of data analysis techniques. Until a completely new perspective on diffusive behavior in crowded microscopic environment is established[3] , our current understanding on diffusion behavior can be in one way or other expressed by the models and fitting techniques that are already in use.

It is perhaps even simpler and better to take a step back and try to understand diffusion decay curves in a single direction than to fit the data obtained from a three dimensional acquisition. Data fitting in single dimension is simpler, gives in to easier interpretation and its validity is easier to check.

The answer to the second question has been relatively unexplored. The dependence of the fitting parameters with variations in the experimental conditions, their stability of the fitting parameters or their instability thereof could inform us on the reliability of the assumptions. The experimental section in this thesis contains a series of measurements performed in this direction.

The last part of this thesis tries to provide an answer to the third question. Based on previous findings, biological understanding of the WM substructures, and the results from the experiments performed during the course of this these, I have discussed the relative importance of the observations and the most likely conclusions that could be drawn. The potential of using this new information in diffusion modeling, its impact in clinic and potential ways to further validate the results have been obtained.

[2]An interesting philosophical discourse on the difference between explaining and understanding can be found in the book from George A. Wright (1971)

[3]Some interesting developments on water mobility near hydrophilic molecules can be found in the review by Gerald Pollack,[151, 152]. An interesting breakthrough research in trapped water has been performed by Naguib et al. [153] They trapped water in 2–5-nm-diameter channels of closed multi-walled carbon nanotubes. Transmission electron microscopy revealed disordered gas/liquid interfaces with greatly retarded water mobility compared to bulk. Future advances on structure of confined might lead way to new ways to look at water diffusion.

# Chapter 5

# Temperature Dependence of WM Diffusion at Very Short Time

*Data! Data! Data! I can't make bricks without clay!* [1]

## 5.1 Introduction

Diffusion-weighted MR (DW-MR) has shown great promise in characterizing WM; nevertheless, understanding of the biophysical characteristics of water diffusion in WM has been elusive. Diffusional decay curves in WM has been shown to be non-monoexponential, and various models have been put forth to explain the relationship between water mobility and the signal obtained [118, 119, 122, 154, 155]. Most models involve cylindrical restriction for intraxonal water in the direction perpendicular to the main fiber orientation and hindrances for extra-axonal water [133, 147, 156]. In a recent paper, Le Bihan [82] has assigned a slowly diffusing pool to the water close to the membranes that form hydration layers. The challenge of effective and realistic modeling is underlined by a recent article [148] comparing 47 different models found in the literature.

Lately, many studies of non-Gaussian diffusion in WM have primarily focused on diffusion properties perpendicular to the fiber directions [140–145]. While methods involving measurements perpendicular to the fiber direction have claimed success in estimating axonal diameters, these studies have implicitly treated layers of myelin as ultimate restrictive barriers and did not consider possible multi-compartmental nature of diffusion data that may be intrinsic to the WM.

An unexplored area related to WM diffusivity is its temperature dependence. For large temperature ranges, the temperature dependence of water diffusivity (in pure water) has been reported to fit a with Speedy-Angell power-law [157]. However, at temperatures near 0 °C, the temperature dependence of water can be approximated by an Arrhenius

[1] Sir Arthur Conan Doyle, The Adventure of the Copper Beeches

equation,

$$D(T) = D_o e^{(-\frac{E_a}{RT})} \tag{5.1}$$

where $E_a$ is the activation energy of apparent diffusion coefficient $D(T)$, $D_0$ is the prefactor, $R$ the universal gas constant. and T the temperature in Kelvin. For bulk water, $E_a$ has been reported to be approximately between 18.8 kJ/mol and 19.2 kJ/mol at room temperature [158, 159].

In tissues containing multiple water pools, measuring diffusion decay against temperature could shed light on the respective environment of the water pools and their interaction. If exchange plays a significant role, according to Kärger's equation [34, 35], the apparent volume fractions of the pools should also change with temperature. The activation energies of different water pools should also provide more insight into the environment they reside in.

In brain WM, layers of compact myelin wrapped by oligodendrocytes around the axon form a unique cellular structure deprived of cytoplasm, in which a small fraction of the water described as 'myelin water' resides in thin layers ($\approx$ 2 nm thick [88]) between apposed membranes held tightly together by a regular pattern of myelin basic proteins (MBP) and proteolipid proteins (PLP). The confinement of myelin water in such narrow environments is expected to result in a depression of the freezing point [160–162].

To better understand the multi-exponential decay of diffusion data, this study measured diffusion with very high b-factors yet at short diffusion time ($\Delta$) and echo time (TE). A short $\Delta$ avoids confounding of the signal with exchange between different pools while a short TE reduces improper estimation of the volume fraction due to potential differences in spin-spin relaxation times ($T_2$) of different diffusing pools. To circumvent restriction effects associated with layers of myelin, measurements were performed parallel to the fiber orientation. Measurements were also repeated at different temperatures allowing the estimation of the activation energies of water diffusion for different water pools.

## 5.2 Methods

**Sample Preparation:** Freshly cut pig brains were acquired from a local slaughterhouse within minutes of sacrifice and immersed in phosphate buffered solution (PBS), adjusted to a pH of 7.4. During transportation, the temperature was maintained between 0 and 4 °C. At the laboratory, a small cylindrical piece of corpus callosum was dissected (4 mm diameter, 4 mm height), excess PBS was lightly soaked up with cotton and the sample was placed in an NMR tube. To maintain humidity, the tube was closed with a small piece of cotton soaked in PBS and sealed with wax. The sample was placed in the tube such that the callosal fibers were aligned with the main magnetic field and

the gradient direction of the spectrometer. All measurements were completed within 9 hours of the slaughter of the animal.

**Data Acquisition:** All NMR data were acquired using a homebuilt FEGRIS-NT 125 MHz spectrometer equipped with a unidirectional ultra-high-intensity magnetic field gradient system that has a peak gradient strength of 35 T/m parallel to the main magnetic field [163, 164]. The temperature of the sample in the bore was controlled using a stream of gaseous nitrogen evaporated from a liquid reservoir and a built-in heater/thermometer regulation system. The initial system temperature was set to 19 ° C. All the pre-measurement steps such as tuning, matching, resonance frequency setting, and determining the optimal pulse durations were performed at this initial temperature.

FID were measured with an initial dead time of 125 $\mu$s and intervals of 5 $\mu$s between successive data points. Following the FID, two sets of Carr-Purcell-Meiboom-Gill (CPMG) [23, 165] measurements were also obtained. The first CPMG set had echo intervals at 0.3 ms, and the second had echo intervals at 2.4 ms. To avoid sample heating due to the radiofrequency pulses, the echo trains were limited to a total of 128 echoes and a repetition time (TR) of 5 s.

Diffusion measurements were performed using a Stejskal-Tanner pulsed gradient spin echo (PGSE) [25] combined with a small readout gradient [164]. Diffusion gradient duration ($\delta$), $\Delta$, TE were kept very short at 0.5 ms, 2 ms, and 4 ms respectively. Gradient pulses were linearly incremented from 0.05 T/m to 31.05 T/m in 32 steps.

The temperature was lowered by 1 °C at each step and the FID was recorded in intervals of 30 s until its signal intensity did not change anymore. Once in equilibrium, an FID and two CPMG decay curves were measured at each temperature. Diffusion measurements were repeated only at 19, 14, 9, 4, 1, −4,−9, and −14 ° C respectively.

**Data Analysis:** Data analysis was performed using a code written in Matlab® (MathWorks, Natick, MA, USA). Since the solid-state NMR signal of ice is not detected in the FID signal at 125 $\mu$s, an average of 4 initial data points from the FID was taken as a proxy for liquid water. Fraction of 'unfrozen' water at freezing transition was calculated as the ratio between signal intensity immediately before and after the freezing transition.

The diffusion data were Fourier transformed to obtain one-dimensional profiles. To avoid influences of sample edges on the echo amplitude, the mean obtained from the central plateau of these one-dimensional profiles were taken as the echo amplitude for each $b$-factor. During post-processing, four data points with maximum $b$-factors were found to be excessively affected by echo shifting and were excluded from the analysis. The system could have corrected for such echo-time shifts [164], but requires twice the number of acquisitions. To avoid measurements at long postmortem delays, the procedure was not followed. The maximum $b$-value used was then 18.8 ms/$\mu$m$^2$.

For temperatures above the freezing transition, both CPMG and diffusion decay curves were obviously non-monoexponential. To identify number of water pools , an inverse

Laplace transform (ILT) was performed for the diffusional decay curves and for the CPMG echo trains with a longer echo interval of 2.4 ms. This involved a regularized non-negative least squares algorithm [166, 167] that minimizes the sum of residuals squared and a regularization function (finite difference approximation of the second derivative of the signal) weighted by a smoothing parameter. For the ILT of CPMG data, a grid of 64 $T_2$ values was logarithmically spaced between 2.4 and 460 ms. For the diffusion data, 24 diffusion coefficients were logarithmically spaced between 0.0053 and 2.71 $\mu m^2/ms$. After ILT, both CPMG and diffusion spectra revealed two major pools.

The data were normalized and fitted to a biexponential decay with a Levenberg-Marquardt non-linear least-squares algorithm that yielded two coefficients and their volume fractions. Diffusion measurements taken close to the freezing transition temperature at $-9$ °C, showed unstable fits and were not further analyzed. To obtain activation energies of the two water pools, natural logarithms of the diffusion coefficients were plotted against inverse temperature and fitted to the Arrhenius equation. The consistency of the results at each temperature suggests that the sample had not undergone any further decomposition during the measurement time.

For the unfrozen water pool that remained 'liquid' below the freezing transition temperature, both CPMG and diffusion data could be sufficiently described with a single exponential decay. Since the relaxation in this pool was very fast, a CPMG train with a shorter echo interval of 0.3 ms was used to estimate $T_2$. Both the relaxation and diffusional decay curves were fitted to a monoexponential decay function using the Levenberg-Marquardt algorithm.

## 5.3 Results

The unfrozen water fraction ranged between 13.5 and 15.0% in all samples (FID data, Fig. **??**). Above the freezing transition, the signal increased slightly with decreasing temperature, reflecting an inverse relationship between equilibrium magnetization and temperature. Below the freezing transition, as more molecules froze with decreasing temperature, the signal kept decreasing steadily. Such information could potentially be used to measure the pore size distribution, but this is beyond the scope of this study.

Immediately after the freezing transition, $T_2$ of the unfrozen water pool had a wide range between 8 and 16 ms for the four samples. A single diffusion measurement for this pool at -14 °C that was fitted to a monoexponential decay resulted in diffusion coefficient ranging between 0.016 and 0.028 $\mu m^2/ms$.

Figure 5.1 is a representative normalized color map of the $T_2$ distribution obtained after ILT of the CPMG decay curves at different temperatures. While the longer $T_2$ component, $T_{2l}$, increased with decreasing temperature, a trend was not observed for the shorter component, $T_{2s}$; probably owing to the difficulty in estimating the latter. The

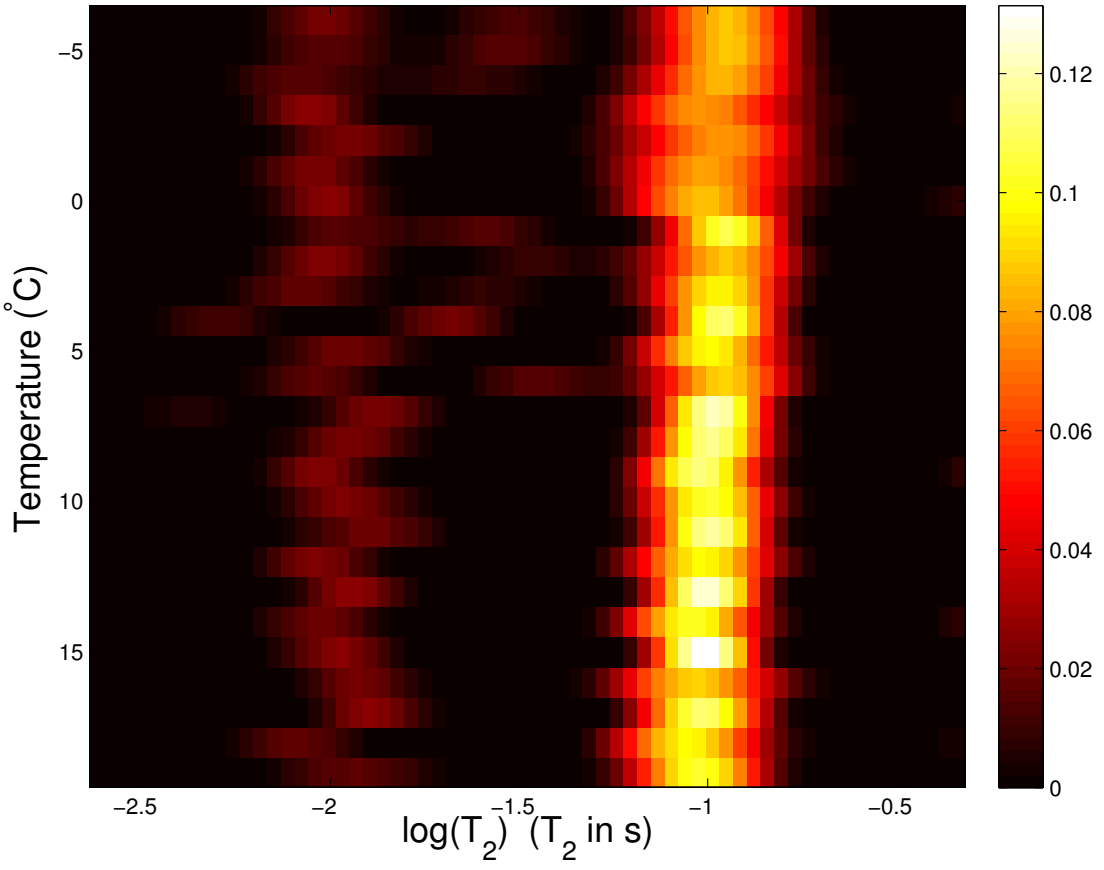



FIGURE 5.1: Distribution of $T_2$ relaxation times for temperatures above the freezing transition obtained after inverse Laplace transform of the individual CPMG decay curves at each temperature. The data is only shown for one sample. While a trend is observed for the long $T_2$ component, no such trend is apparent for the short $T_2$ component.

$T_2$ time constants deduced from biexponential fits also retained this trend. Figure 5.2 shows the temperature dependence of the $T_2$ values for the two components obtained with biexponential fits. The volume fraction for $T_{2s}$ was estimated to be $19.2 \pm 0.1\%$ for all samples and all temperatures before the freezing transition.

Inverse Laplace transform of the diffusion data also revealed bi-modal peaks that suggested two major components. Figure 5.3 shows distribution of diffusion coefficients at different temperatures for all four samples. Both the faster diffusing component, $D_f$, and the slower diffusing component, $D_s$, decreased with temperature. A biexponential fit, consistent with two water pools with different diffusion characteristics was sufficient to model the data. Assuming negligible exchange between the two pools, at 19 °C and 2 ms diffusion time, the fast pool is found to have a diffusion coefficient of 0.59 $\mu$m$^2$/ms corresponding to a diffusion length, $l_D \equiv \sqrt{2D\Delta}$, of around 1.54 $\mu$m and for the slower pool, a diffusion coefficient of 0.05 $\mu$m$^2$/ms had a diffusion length of 0.45 $\mu$m.

Figure 5.4 is the Arrhenius plot of the two diffusing pools. The activation energies

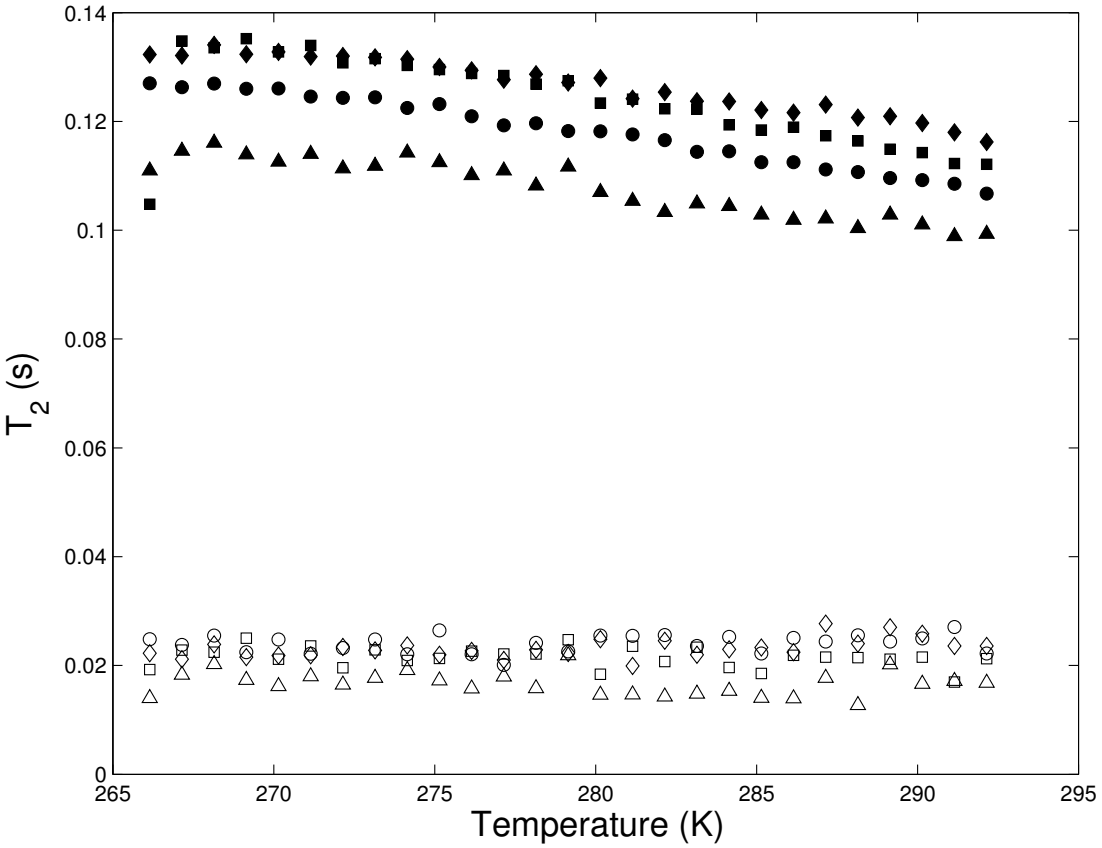



FIGURE 5.2: $T_2$ relaxation times against temperature for the short (open symbols) and long (solid symbols) components obtained from a biexponential fit of the CPMG decay curves. Different symbols represent measurements in different samples.

| | **19 °C** | **14 °C** | **9 °C** | **4 °C** | **-1 °C** | **-4 °C** | $\mathbf{E}_a$ |
|---|---|---|---|---|---|---|---|
| $\mathbf{D}_f$ | $61 \pm 5$ | $54 \pm 3$ | $50 \pm 4$ | $45 \pm 3$ | $42 \pm 3$ | $37 \pm 3$ | 13.9 |
| $\mathbf{D}_s$ | $5 \pm 0.3$ | $5 \pm 0.6$ | $5 \pm 0.2$ | $4 \pm 0.4$ | $4 \pm 0.4$ | $4 \pm 0.4$ | 8.3 |
| $\mathbf{V}_f$ | $54 \pm 4$ | $53 \pm 5$ | $52 \pm 4$ | $52 \pm 4$ | $52 \pm 4$ | $51 \pm 3$ | $--$ |

TABLE 5.1: ADC at different temperatures obtained from all the samples. $D_f$ and $D_s$ are the apparent diffusion coefficients of the fast and the slow components. $V_f$ is the volume fraction ( in %) of the fast component. The last column shows the activation energy of diffusion for the two pools. The units are in $10^{-2}\,\mu$m$^2$/ms for the diffusion coefficients and in kJ/mol for the activation energy.

for fast and slow diffusion pools were found to be 13.9 and 8.3 kJ/mol respectively. Extrapolation of the results for the two pools to human body temperature 37 °C produced diffusion coefficients of 0.84 $\mu$m$^2$/ms for the fast pool and 0.06 $\mu$m$^2$/ms for the slow pool. Table 5.1 summarizes the results from the diffusion measurements.

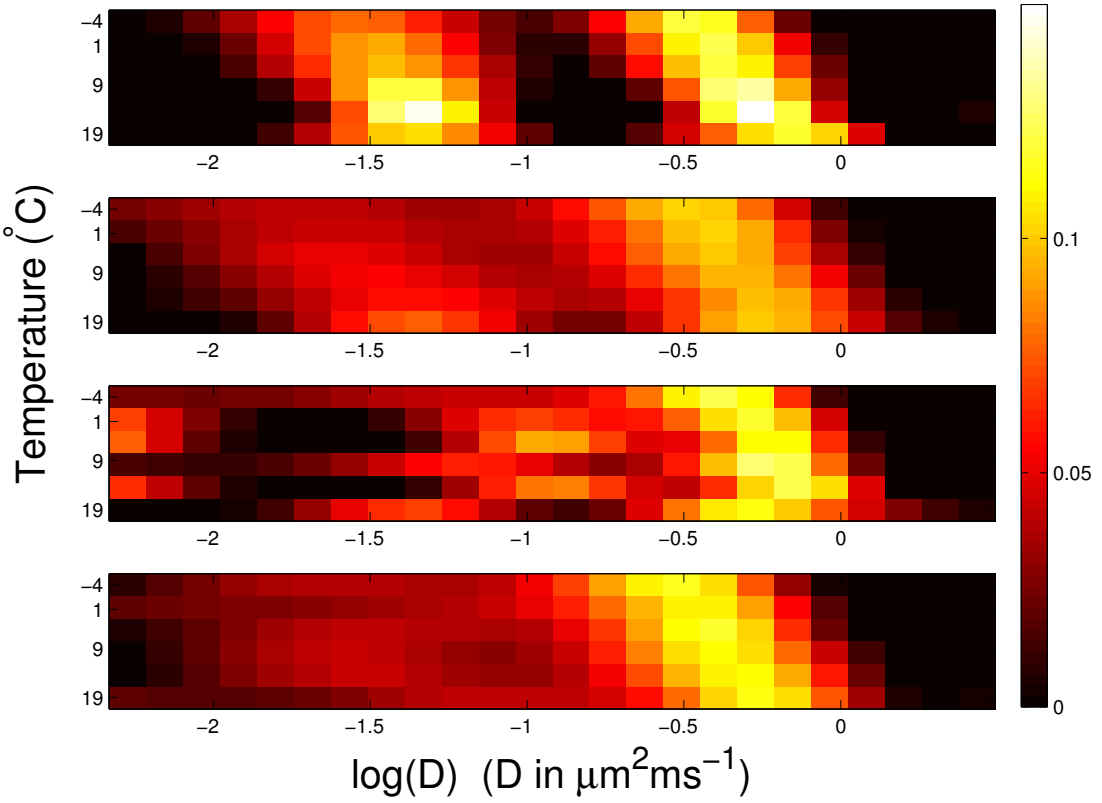



FIGURE 5.3: Distribution of diffusion coefficients obtained by inverse Laplace transform of the diffusional decay curves for all four samples. Both the fast and slow pool show increased diffusion with increasing temperature. In this figure, for consistency, data from all the four samples and different temperatures were analyzed using same regularizer. This has caused some discrepancy in sample 3 for some temperatures. Standard ILT practices tend to choose the appropriate regularizer for each data set.

## 5.4 Discussion

The thin layers of myelin water trapped in myelin lamellae are typically in the order of a few nm. It is likely that water molecules trapped between these layers experiences a freezing point depression [168]. Most of the signal remaining after the freezing transition can be attributed to the water residing between layers of myelin. The unfrozen water's volume fraction around 14% and its $T_2$ around 12 ms is close to values estimated for the shorter $T_2$ component before the freezing transition. The shorter $T_2$ component was found to be around 20 ms with a volume fraction of about 20%. This small water fraction with short relaxation times has been widely considered to be myelin water fraction [169, 170]. Close match between the unfrozen water fraction and the shorter $T_2$ pool strongly suggests that most of the short $T_2$ signal arises from water inside the myelin layers. To our knowledge this is the first time that the contribution of myelinated layers to the shorter $T_2$ have been established using a complementary physical technique.

While the shorter $T_2$ component attributed to myelin water fraction showed very little

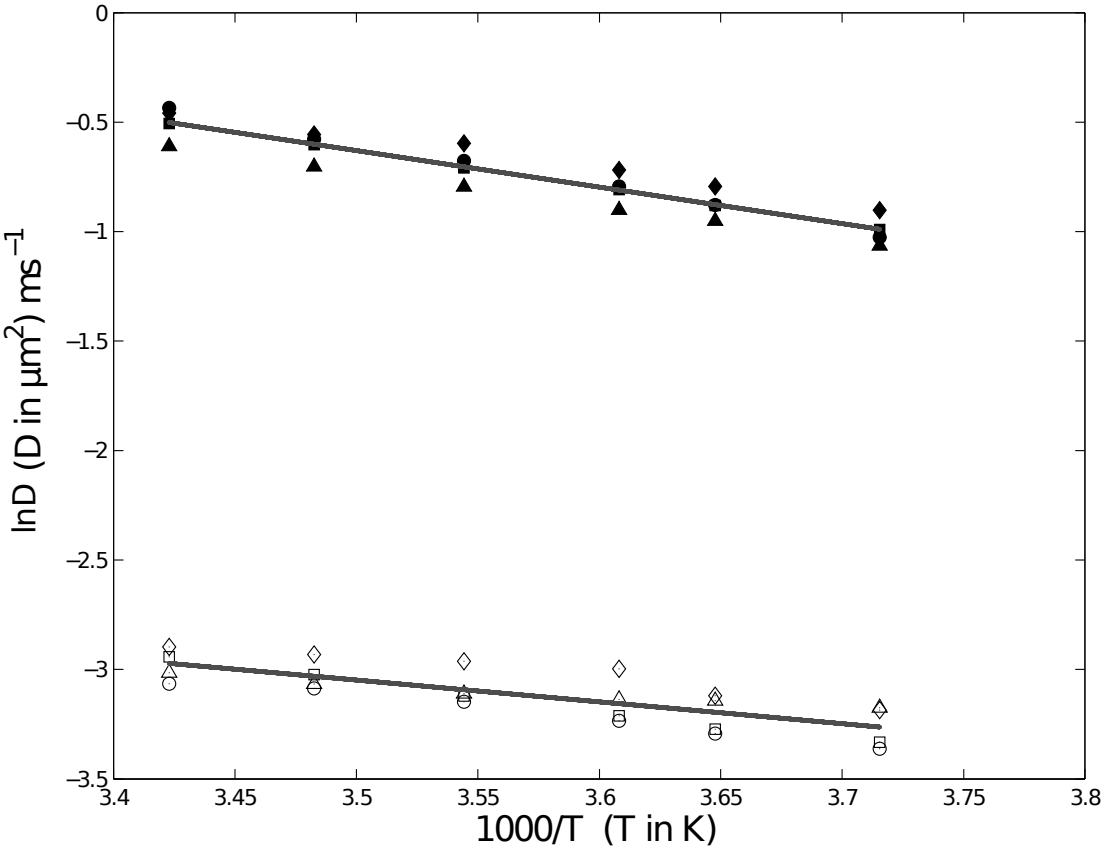



FIGURE 5.4: Arrhenius plot: Plot of the natural log of the fast (filled symbols) and slow (open symbols) diffusion coefficients against inverse temperature. Different symbols represent measurements in different samples. The lines are linear fits for the two components. Activation energies of diffusion were estimated to be 13.86 kJ/mol and 8.26 kJ/mol for the fast and the slow pool respectively.

dependence on temperature, the longer $T_2$ seems to decrease with increasing temperature. In muscles, a similar weak negative dependence has been previously attributed to hydrogen exchange between water and proteins [171]. A weakly negative dependence of $T_2$ on temperature could also arise from the molecular motion in a heterogeneous medium. Increased diffusion at higher temperatures could have resulted in stronger contribution from internal gradients to reduced measured $T_2$.

Due to the short TE of 4 ms employed, myelin water also contributed to the PGSE signal. For $T_2$ of 20 ms, myelin water still retains 80% of its signal at this echo time contribution 16% to the total PGSE signal. The ADC of the unfrozen myelin water at -14 °C was found to be around 0.025 $\mu$m$^2$/ms suggesting that the myelin water contributes to the slow water pool.

The water dynamics is largely determined by hydrogen bond networks. Most of the myelin water is hydration water, but are relatively isolated from water in the rest of the WM [172]. Ordering of water is to be expected in water molecules close to a phospholipids which are contained in the myelin sheath [52]Water molecules close to hydration

layers of the membranes are known to have retarded translation properties compared to water in bulk [173]. Furthermore, the presence of MBP and PLP also provide steric hindrances to water diffusion pathways. It has also been shown that water molecules bound to the bilayers have a decreased hydrogen bonding strength [52]. Magnetic field dependent $T_1$ study of unfrozen water fraction in mouse tissue also reported a decrease in activation energy for bound water [174]. The activation energy of the slow pool is close to the average energy required to break a hydrogen bond in a locally structured domain [175] which is similar to the latent heat of melting that breaks about 10% of the hydrogen bonds [176]. A drop in activation energy suggests some sort of local ordering of the water molecules. In light of these results, we maintain that the myelin water is part of slow diffusing pool and has a lowered activation due to its association to the hydration layers of the membranes (or compacting proteins MPB and PLP) . A recent explanation by Le Bihan [82] assigns fast and slow diffusion pools to bulk water, and water found near the surface of the membranes respectively.

Myelin water contribution to the PGSE signal is only around 16% ( $T2$ = 20 ms, $TE$ = 4 ms, Myelin Fraction = 20%). It does not account for all of the slow water pool found in the diffusion decay curves. Many studies, performed with longer echo times and thus very little myelin water contribution, have also reported similar multi-exponential decay [154, 155, 177]. ILT showed that the distribution of the ADCs was clearly bimodal. This is a strong indication that the diffusion decay curve did not quite originate from water molecules residing in a distribution of similar environments (as would be when they come from cell of different sizes).

Restrictions, in principle, could give rise to a non-monoexponential signal decay that need not necessarily resemble a biexponential decay. Restriction effects can be conveniently described in terms of free diffusion regime, motional narrowing regime, and localization regime depending on the dephasing length, the diffusion length and the average distance between restrictive barriers [40, 42, 43]. In the free diffusion regime, diffusing particles are not to experience boundary effects; in this case, observation of a biexponential decay simply suggests the existence of two water pools with differing diffusivities. In the motional averaging regime, boundaries are considered to be so close to each other that an increase in true diffusivity due to an increase in temperature leads to an apparent decrease in the measured diffusion coefficient [40]. For our measurements, the motional narrowing regime would translate to increasing diffusion coefficients with decreasing temperature. We can thus effectively rule out the motional narrowing regime.

This leaves us with the localization regime. In this regime, Hürlimann et. al [43] for one-dimensional parallel plates and by Milne and Conradi [178] for cylindrical tubes, showed deviations from monoexponential decay. As discussed in section 3.6.3 this occurs when the dephasing length $(D_0/\gamma g)^{1/3}$ is smaller than the diffusion length $\sqrt{D_0\delta}$. With decreasing temperature, intrinsic diffusivity of the water molecules decreases, thus

changing the ratio between the dephasing length and the diffusion length. Such change should have shown a corresponding changing in volume fractions with temperature [179]. However, this was not observed. While some interesting combinations of geometry and pore size might be proposed to explain our data, it is generally accepted that diffusion of water parallel to the fiber direction remains unchanged for long diffusion times. Therefore, it is unlikely that a restriction effect alone could account for the deviation from monoexponential decay observed in our measurements.

Axons, oligodendrocytes and astrocytes are proliferated with many microtubules, neurofilaments and microfilaments. Microtubules are essentially cylindrical tubes that are about 25 nm in diameter, with a packing density of around $70/\,\mu\text{m}^2$ and can extend up to hundreds of micrometers in length, in a dynamic process of growth and shrinkage used for motion along the axon [180, 181]. These tubular structures are all aligned parallel to the axons primary axis and are highly anisotropic. Almost a third of the water population inside the microtubules is hydration water. Furthermore, recent studies show that the lumen of the axonal microtubules are filled with particles about 6 nm in size and separated at an average distance of 15 nm [78, 79]. The microtubules are very anisotropic: at the timescales probed by our diffusion times, exchange between water molecules in these structures and the water outside the tubules is limited. The water in these tubules are not bound water like in myelin - the rotational mobility and the corresponding $T_2$ could be similar to water elsewhere, but the presence of large fraction of hydration water and hindering particles could drastically reduce the mobility of water inside the tubular lumen. The water in the tubules could exhibit a slower diffusion that is grouped to the diffusion of myelin water while their $T_2$ is not distinguishable from the fast diffusing pool. It has also been reported that a decrease in microtubules due to introduction of methylmercury accompanies an increases in the diffusion of water along the fiber direction [77]. Microtubules in the axons are oriented along the fiber axis, along direction of the processes in glial processes and randomly in the glial cell body. Microtubules perhaps contribute around 5% to the overall signal and water trapped in tubular bodies and cell organelles could perhaps account for some 10% of the signal, but this clearly is not enough to account for the volume fraction of slow water found in this study.

A major fraction of the WM volume consists of glial cells and their processes. The glial cell bodies are quite large and the water molecules in there should show more or less fast isotropic diffusion. The glial processes are more or less randomly oriented and occupy a significant volume fraction of the WM. A recent study in optic nerves of guinea pigs report that about 20% of the volume is occupied by the glial processes [83]. The cytoskeleton of these processes is also aligned with the process orientation. Like axons, water diffusion in these processes should be anisotropic. At such short times probed by this study, very little exchange between the processes and the outside can

be expected. The diffusion decay curves in such isotropically distributed anisotropic tubes is also non-monoexponential [97] and can very well be fitted by a biexponential. As seen in Eq. 4.2, if ADC is to be calculated from high b-values, the estimated ADC is b-value dependent giving ADC(b)= $D_\perp + 1/2b$. Assuming that the diffusivity along the processes is same as the diffusivity of intraaxonal water (0.6 $\mu$m$^2$/ms), the signal from all b-values greater than 8 ms/$\mu$m$^2$ will contribute to the non-monoexponential tail. Since the processes are mostly narrower than 0.5$\mu m$ [98], $D_\perp$ can be ignored. In this case the ADC of the 'slow' component will be between $1/(2 \times 8)$ and $1/(2 \times 18.8)$. The slow component measured in this study lies well within this range. An interesting aspect about this pseudo 'slow' component due to the three dimensional distribution of glial processes is that it depends only on the b-value and is therefore independent of the temperature. This would effectively result in a zero activation energy. Inclusion of this factor in the slow component has perhaps underestimated the activation energy for other components that also contributed to the slow pool.

## 5.5 Conclusion

Our observation at a short diffusion time of 2 ms provide clear evidence of two distinct pools of mobile water within ex-vivo WM, with activation energies differing by 50%, and volume fractions that are independent of temperature. This implies that the non-monoexponential diffusion signal indeed arises from separate water compartments, and is not merely an effect of restriction within a single compartment by cell membranes.

Models that attribute non-monoexponentiality to restrictions assume that all water molecules reside in more or less similar environments. Distinction should be made between restriction and hindrances: restricted water is confined within a structure whereas for hindered water boundaries only contribute to effect an overall reduction in mobility. In all biological tissues, water diffuses more slowly than in pure water, and can be considered hindered. As in all studies using biological tissues, both the fast and slow diffusion pools observed in this study were definitely hindered.

Biexponential fits of the diffusion data revealed two diffusion coefficients that were an order of magnitude apart. The activation energies of the fast and slow pools were different from each other by as much as 50%. However, the volume fractions of the two pools did not depend on temperatures. Invariance of the volume fractions on temperature implied minimal exchange between the two water pools. The observation of two distinct pools of mobile water with different activation energies, but consistent volume fractions at different temperature suggests that the non-monoexponential signal decay to originate from water pools experiencing different environments.

Myelin water is readily distinguished due to its short $T_2$. It also diffusion slowly. But, there seems to be another pool of water that has slow diffusion properties but its $T_2$ is

not distinguishable from the fast pool. Together with values reported from the literature our results suggest that the slow diffusion pool arises from two water pools (i) water molecules that are in exchange with the water in the layers of hydrated water close to the cytoskeletal structures like microtubules and in the membranes like myelin sheath and (ii) a non-monoexponential component due to three dimensional distribution of glial processes. The rest of the axonal water and water inside glial cell bodies contribute to the fast pool.

## Note

It would have been nice to be able to perform similar measurements perpendicular to the fiber direction. However, when the sample was aligned perpendicular to the main magnetic field, the difference between the long and short time CPMG train was significant, suggesting strong internal gradients due to susceptibility differences within the sample. A strong TE dependence could also be seen in the diffusion decay curve at high b-values. Bipolar PGSE sequence with gradients of alternating polarity could be used to mitigate these effects but that did not seems to simplify the understanding of the decay curves.

# Chapter 6

# Diffusion Measurements *in vivo*

## Introduction

Self-diffusion of water in the WM varies considerably depending on its orientation [101], location [182], maturity [183–185], pathology [186–188], and death [189]. However, limited SNR along with a poor understanding of the underlying origin of non-monoexponential diffusion decay has severely handicapped high b-factor DWI and rendered it as an unfavored probing mechanism in clinic as well as in research.

A good *in vivo* imaging needs to provide adequate spatial resolution and also higher SNR. The human brain is extremely compact. Structures that are functionally and anatomically distinct are present within close proximity. High spatial resolution is a necessity to distinguish between those sub-structures. Sensitive models also need high SNR to obtain robust parameters. A high field strength, along with stronger field gradients and better RF coils help to achieve these requirements. The physical basis of non-monoexponential diffusional decay in human brain is not yet fully understood; availability of reliable, high resolution, *in vivo* data would be a major boost for such an endeavor.

This chapter details the experiments performed in human brain *in vivo*. The edge here rests primarily upon the availability of a 7 Tesla whole body human scanner equipped with a 70 mT/m gradient system. This is a huge advantage compared to a 3 T system with 40 mT/m gradients. Although the basics of DWI remains the same, general technical and safety considerations vary considerably at higher field strengths.

The first section of this chapter briefly outlines different issues that need to be addressed at high fields and briefly explains the experimental setup. The second part is a compilation of reports on different diffusion measurements mostly performed in a chronological order. Any issues that had not been necessary to consider during the initial experimental design but was later addressed is described as it appears and is implemented for subsequent measurements.

## 6.1 Experimental Design

The increase in field strength of clinical scanners and the corresponding improvement in SNR have led to considerable advances in the quality and spatial resolution of *in vivo* DWI of the brain. DWI provides quantitative maps of ADC and reveals orientational structure in anisotropic tissue such as brain WM. At high b-factors diffusional decay curves deviate from monoexponential behavior [120, 122, 184] and promise to provide further information on the tissue micro-structure [132, 190]. DWI at high b-factor has also been reported to be sensitive in characterizing pathological conditions such as brain tumors and Alzheimer's disease [187, 188].

For high b-factor DWI, application of diffusion gradients selectively attenuates signal intensity, and much of the useful orientational information arises from data with high diffusion gradients. However, due to exponential nature of diffusional decay curves, high b-factor DWI intrinsically requires high SNR and consequently its applications have been relatively rare. Nonetheless, to understand non-monoexponential nature of diffusional decay of MRI signal in brain tissue and to benefit from its application is a long-sought-after goal. Given the current technological state, a prudent choice of parameters for acquisition and data analysis would assist in achieving this goal. DWI should benefit from the higher SNR available at ultra- high fields, however, higher fields are also associated with higher inhomogeneities, high specific absorption rate (SAR) and shorter $T_2$ relaxation that need to be properly considered before designing any such measurement protocol.

The most common DWI sequence for clinical scans is the twice-refocused spin echo EPI (TrSE-EPI) developed by Reese et al. [191]. TrSE-EPI has been very widely used for its ability to compensate intrinsically for the effects of eddy currents. This technique, first proposed in the context of perfusion imaging by Feinberg and Jakab [192], employs four gradient lobes such that at the end of the fourth lobe the eddy currents generated by each of these gradients cancel each other. However, TrSE-EPI employs two 180° refocusing pulses, which at ultra-high fields contribute significantly to RF power deposition (SAR) and seriously limit *in vivo* measurements. Applying two refocusing pulses also exacerbates the non-uniformity in the image intensity related to the inhomogeneity in the RF field.

At higher field strengths, the short $T_2$ in brain WM is also a major concern for DWI [193, 194]. Application of the long duration diffusion sensitizing gradients needed to achieve high b-factor DWI results in a major loss of SNR to $T_2$ relaxation. Gain in intrinsic SNR from using a 7 Tesla system as compared to a 3 Tesla system with otherwise identical noise level is approximately 7/3. Any spin echo based imaging sequence

would therefore benefit at 7 T only at short TE times where

$$TE < \ln\left(\frac{3}{7}\right)\frac{T_2(7T) \times T_2(3T)}{T_2(7T) - T_2(3T)} \tag{6.1}$$

Given that the $T_2$ for brain white matter is around 90 ms at 3 T and 50 ms at 7 T, 7 T can only maintain its advantage when TE is shorter than 90 ms. For high b-factor spin-echo DWI acquisitions with long echo times (TE) signal loss due to the faster $T_2$ relaxation at higher field strengths outweighs the signal gain due to higher intrinsic SNR.

Fortunately, the spin lattice relaxation ($T_1$) in brain tissue increases with increasing field strength. The decreased $T_2$ along with an increased $T_1$ suggests that using stimulated echoes for diffusion preparation could be well suited for high field, high b-factor DWI [195].

A stimulated echo requires three radio-frequency (RF) 90° pulses - the signal arises from one-half of the diffusion-labeled magnetization that is stored in the longitudinal axis during the time interval $TM$ between the second and third pulse. During $TM$, the magnetization undergoes longitudinal ($T_1$) relaxation, allowing long diffusion times and high b-values to be achieved with minimal signal loss to $T_2$ relaxation. The stimulated echo has long been used in MR imaging to exploit the T1 dependence of the signal. Since the diffusion-labeled magnetization is stored in the z-direction during the time TM, it undergoes slow longitudinal T1 relaxation, allowing very long diffusion times and high b-values to be achieved with minimal signal loss due to transverse relaxation. Although STEAM utilizes only half of the intrinsically available signal, for long $\Delta$ measurements stimulated echo setup is more effective than losing most of the signal to rapid transverse relaxation. When $T_1 >> T_2$, inspire losing half the magnetization, use of longer diffusion times to achieve the desired b-value is more effective than losing most of the signal to rapid transverse relaxation.

Diffusion weighted stimulated echo acquisition mode (DW-STEAM) also allows measurements at very long diffusion times. As mentioned before, compared to the spectrometer used in *ex vivo* study, human scanners are equipped with very weak gradients and therefore, for *in vivo* studies pushing the limits with long diffusion time measurements could allow valuable insights.

For fast data acquisition, EPI was chosen as the preferred k-space sampling method. High field strength, high b-factor, long diffusion time DWI was achieved using diffusion weighted stimulated echo acquisition mode echo-planar imaging (DW-STEAM-EPI).

### 6.1.1 STEAM sequence

Fig .6.1 shows standard schematic for stimulated echo diffusion weighted acquisition. The first row shows three RF pulses needed for a stimulated echo formation. The second row shows slice selection gradients and their rephasing gradients for the three RF pulses

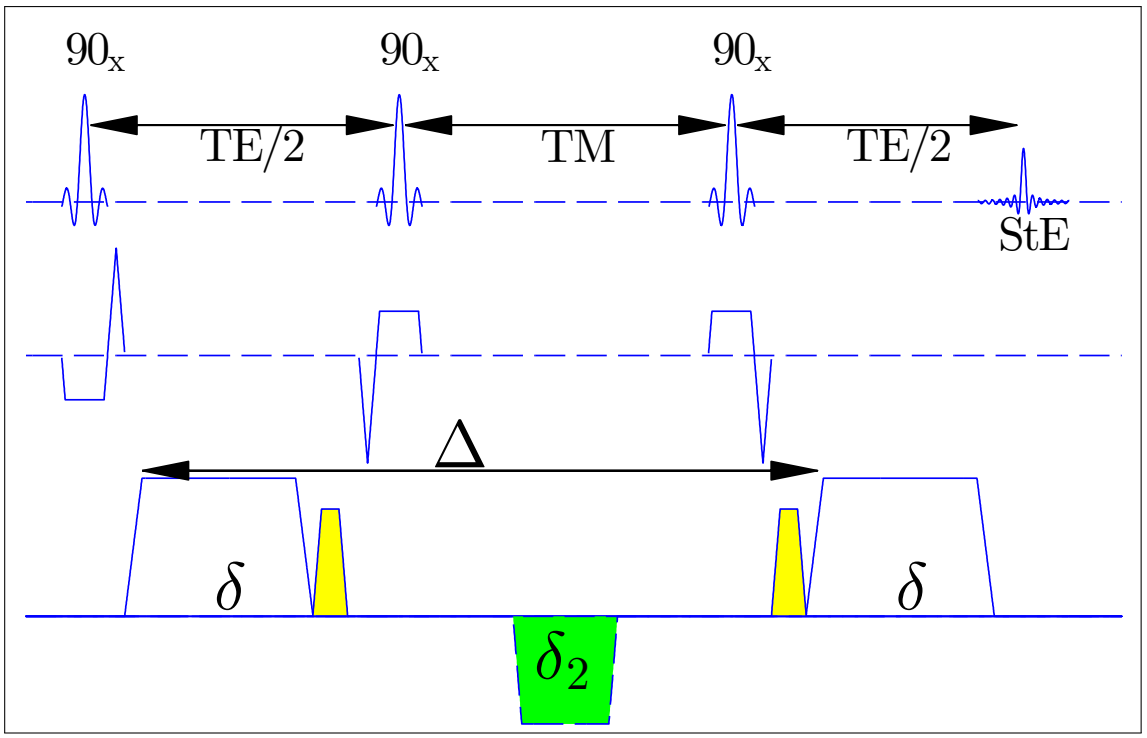



FIGURE 6.1: Schematic of STEAM pulse sequence. The first line shows RF pulse, the signal and echo time. The second line shows the gradients for the slice selection. The third line shows the diffusion gradient length $\delta$ during the $T_2$ decay time and the homogeneity spoiling gradient ($\delta_2$) during $T_1$ decay time. To save on TE and signal loss due to $T_2$ relaxation, when the diffusion gradients are strong, the crusher gradients (shaded yellow) are redundant and can be removed.

and the third row shows the diffusion gradients along with the crusher gradients (shaded yellow) and the homogeneity spoiling gradient (shaded green). The stimulated echo is formed when the second $90_x$ pulse applied after time TE/2 rotates the y-component of the magnetization into z-axis and is subsequently recalled using the third pulse. The echo is formed at time TE/2 after the third pulse. Between second and third pulse, the spins are in the longitudinally polarized state and only undergo $T_1$ relaxation. However, since the second $90_x$ pulse leaves x-magnetization in the transverse plane, the stimulated echo accounts for one-half of the total magnetization. The remaining x-component can give rise to spin echoes; primary echo (between first and second pulse) and secondary echo (between second and third pulse) that can interfere with the stimulated echo. The homogeneity-spoiling gradients applied between second and third pulse (shaded green) can kill the unwanted transverse magnetization while the longitudinal magnetization remains unaffected. Similarly crusher gradients (shaded yellow) are used to kill unwanted interference from the FID due to the third $90_x$ pulse.

#### 6.1.1.1 Modification to the STEAM-EPI

The following modifications were performed in STEAM-EPI procedure to fit measurement in 7 T system.

**Crusher Gradients** The minimum gradient integral for crusher gradient is determined by the length of the acquisition volume in the direction of the crusher gradients. It is required that no signal should remain if the data is acquired immediately after the crusher. The phase difference, $\Delta\phi$, between the initial and final position of the spins is

$$\Delta\phi = (\gamma\, G\, t_c)\, l \tag{6.2}$$

where $(\gamma\, G\, t_c)$ is the gradient integral, $l$ is the length of the acquisition volume, $\gamma$ is the gyromagnetic ratio, $G$ is the gradient amplitude and $t_c$ is the effective duration of the crusher gradients. In a 2D slice-by-slice acquisition, the effective length in read or phase direction is around 20 cm (size or human head) compared to 2 mm in the slice direction (slice thickness). To dephase signal from a 2 mm length it is enough to have a crusher gradient integral of $4000\pi\, Hz$ but only 100 times smaller for 20 cm. When $\Delta$ is 2 s it translates to a b-value of 0.3 ms/$\mu$m$^2$. This requires stimulated echo diffusion acquisition to have a minimum b-value that is proportional to $\Delta$ employed. If the crusher is applied in read or phase direction, the same effect can be achieved with very small contribution to the diffusion weighting.

In the presence of diffusion gradients that satisfy the required gradient integral criteria, the crusher gradients are redundant and can be done away with. This principle has been applied in sequence design, the crusher gradients and the diffusion gradients have been combined into one making sure that the minimum b-value requirement is always satisfied.

**Slice Selection Gradients** The slice selection gradients are useful for imaging and can be employed in any direction to effectively select a volume of interest. Turner et al. used STEAM-EPI for inner volume imaging to improve resolution and homogeneity in an EPI slice [196]. Fig. 6.1 shows that the slice selection gradient for the first RF pulse has its polarity reversed relative to the later pulses. This is akin to the slice selective gradient reversal method that is used to suppress unwanted chemical species in spin echoes [197, 198]. Reversing polarity between gradient pulses ensures that the through-plane chemical shift is in opposite directions for the two pulses so that the spatial profile of the two pulses are the mirror image of each other around the measurement slice. The unwanted chemical species that have a slight chemical shift relative to water proton resonance frequency only receive either of the RF pulses and cannot follow the same coherence pathway as the water protons. For *in vivo* brain imaging this technique is mostly used to suppress fat with a chemical offset of 2.2 ppm. Fat suppression using

gradient reversal can save us from using spectrally selective RF pulse as a preparatory step for fat suppression. For *in vivo* imaging at high fields, spectrally selective pulses impose additional SAR load and are better to be done away with.

**Eddy current compensation** Application of strong gradient also produce a lagging eddy current response with opposite polarity to the applied gradients. Considering eddy current decay as exponential and isotropic for gradients in all directions, we can write the residual eddy current after the final gradient waveform as

$$I_e = I_0 \left(1 - e^{-\lambda\delta} - e^{-\lambda(\delta+d)} + e^{-\lambda(\delta+d+\delta_2)} + e^{-\lambda\Delta} - e^{-\lambda(\Delta+\delta)}\right) \tag{6.3}$$

where $I_e$ is the final eddy current, $\delta$ is the duration of the diffusion weighting gradients, $\lambda$ is the eddy current decay constant and $d$ is the distance from the end of the homogeneity spoiling gradient to the beginning of second diffusion gradient that can be fixed beforehand. $I_e$ is effectively nulled if

$$\delta_2 = \frac{1}{\lambda} ln \left[\frac{e^{-\lambda(\delta+d)}}{e^{-\lambda d} + e^{-\lambda(\delta+d)} - e^{-\lambda\Delta} + e^{-\lambda(\Delta+\delta)}]}\right] \tag{6.4}$$

For practical values of $\lambda$, each $\delta$ is associated with a unique $\delta_1$. As seen from the sequence diagram, the eddy current compensating gradient is played out during the TM time for the stimulated echo, and thus does not affect phases of the spins forming the stimulated echo. It therefore does not contribute towards diffusion weighting. Similar eddy current compensation method for stimulated echo has been previously used by Steidle and Schick [199] where they split the spoiler gradient into two parts. In phantom, the best images were obtained when $\lambda = 32$ ms was used. However, the difference between different values of $\lambda$ was quite subtle.

**B-value** The gradient scheme used for diffusion weighting in STEAM-EPI is similar to that used in Stejskal-Tanner spin echo preparation: gradient duration ($\delta$) and the diffusion time ($\Delta$) can be precisely defined. Considering trapezoidal gradients, we have

$$\Delta = \delta + \epsilon + TM + Rt \tag{6.5}$$

where $\epsilon$ is the ramp time for the trapezoidal gradients and $Rt$ is the time required for the RF pulse [200]. The b-value can be defined as

$$b = \gamma^2 G^2 \left\{\frac{2\delta^3}{3} + \delta^2(TM) + \delta^2(Rt) + \ \delta^2\epsilon - \frac{\delta\epsilon^2}{6} + \frac{\epsilon^3}{30}\right\} \tag{6.6}$$

### 6.1.2 Data Acquisition

The RF coil used for signal detection is a 24 channel phased array receive coil with volume excitation (Nova Medical, Wilmington, MA, USA). Multichannel data from the phased array coils were reconstructed with adaptive combine [201] approach so that the resulting magnitude image has a Rician distribution [202]. The weights for the linear combination in each pixel are derived from the sensitivity of each coil in that pixel and the noise correlation matrix. As weighted sum of Gaussians, the combined complex image still has a Gaussian noise distribution and the resulting magnitude image has a Rician distribution [202].

### 6.1.3 Measurement Procedure

All studies were carried out with the approval of the ethics committee of the University of Leipzig, Medical School. Informed consent was obtained from each volunteer before scanning.

All the measurements reported in the following sections have been performed on a 7 T whole-body MR scanner (MAGNETOM, Siemens Healthcare, Erlangen, Germany) housed in the Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig. This magnet is equipped with a Siemens-AC072 whole body gradient coils that can achieve a peak gradient strength of 70 mT/m in each direction (x, y, z). The maximum slew rate of the gradient system is 200 T/m/s. When necessary a combination of two orthogonal directions can boost effective gradient strength by a factor of $\sqrt{2}$ but care should be taken about the duty cycle. The gradient safety watchdog is quite sensitive to heating of these gradients. For poorly prepared DWI protocol, it is not uncommon to find a sequence stop suddenly halfway through the measurement process.

### 6.1.4 Selecting WM voxels with coherent fiber bundles

To select WM voxels with only single fiber bundles, we calculated coefficients of linearity $\left(c_l = \frac{\lambda_1 - \lambda_2}{\lambda_1}\right)$, planarity $\left(c_p = \frac{\lambda_2 - \lambda_3}{\lambda_1}\right)$ and sphericity $\left(c_s = \frac{\lambda_3}{\lambda_1}\right)$ where $\lambda_1$, $\lambda_2$, and $\lambda_3$ are the first, second, and third eigenvalues of the diffusion tensor estimated at that voxel. These coefficients describe proximity of the tensor to a line, plane and sphere [103]. The voxels selected as a single fiber bundles had to fulfill $FA > 0.5$ for all diffusion times and $\overline{c_l} \geq 0.4$, $\overline{c_p} \leq 0.2$, $\overline{c_s} \leq 0.35$, where $\overline{c_l}$, $\overline{c_p}$,$\overline{c_s}$ are the mean values obtained for all diffusion times. This method has been previously used by Fieremans et al. to select single WM bundles in vivo [203]. After FA thresholding only voxels present were those in the corpus callosum, cerebrospinal tract and cingulum that are the three major white matter fiber bundles. Only these voxels were further examined.

For all measurements, it has been tried to keep the imaging protocol as consistent as possible. Some adjustments in field of view or the imaging matrix had to be done

depending on the head size of the volunteer or tilt in the scanner. The scanning slice was normally tilted close to 15 ° to the axial position. This slight tilt has been found very effective in avoiding susceptibility related artifacts due to air in paranasal sinuses. The acquisition bandwidth was 1930 Hz/pixel and the k-space was partially sampled at 5/8. Owing to complex noise behavior, parallel imaging technique was not applied. Images were acquired with 2 mm isotropic resolution using 2D single shot DW-STEAM-EPI.

All the following measurements were focused to understand DW data in WM voxels with single coherent fiber bundles. The measurement techniques could be applied for any other structure of interest.

## 6.2 DWI at low b-values: DTI at variable diffusion times

### 6.2.1 Introduction

DWI at low b-values ( $\sim$1 ms/$\mu$m$^2$ ) evaluates the ADC in a given voxel. These low b-value *in vivo* diffusion measurements have generally used diffusion times ($\Delta$) between 10 and 100 ms. To our knowledge, Le Bihan et al. [112] were the first to report about diffusion time dependence in human brain white matter *in vivo*. They reported no change in diffusion coefficient for measured diffusion times between 16 and 79 ms. The ADC was reported to be about 1 $\mu$m$^2$/ms and 0.45 $\mu$m$^2$/ms for diffusion parallel and perpendicular to the white mater fiber orientation respectively. Shortly thereafter, another study by Horsfield et al. [204] reported ADC to be time dependent for diffusion times ranging between 50 and 800 ms. At diffusion time of 50 ms, diffusion coefficients in the corpus callosum was reported to be about 1.8 $\mu$m$^2$/ms parallel to the axonal axis and 1 $\mu$m$^2$/ms radially. In both these studies, white matter fiber orientation was assumed to be known from anatomy.

Low b-value data have been extensively used in diffusion tensor imaging (DTI). DTI can account for a single fiber orientation and fails to describe signal from WM where multiple fiber orientations are present in a single voxel. Notwithstanding these limitations, DTI can still help us reveal gross fiber orientations in WM regions where the fiber bundles are uniformly oriented. Clark et al. [205] used 6 directions to evaluate diffusion tensor between 8 and 80 ms and found the ADC and FA in all directions to be independent of the diffusion time. Using realistic simulations, Ford et al. [206, 207] showed that time dependence of ADC is largely dependent on the extracellular space and permeability. They also reported that the diffusion coefficient transverse to the WM fibers could attain its asymptotic value within one millisecond [207]. More recent studies on mice in vivo [208] and monkey brain *ex vivo* [209] have reported time dependence of DTI parameters between 30 and 280 ms.

In heterogeneous media, the study of time-dependent ADC and its relation to the underlying micro-structure is an active area of research. In heterogeneous media, the central portion of the PDF appears Gaussian while the tail could take other forms [210]. The central limit theorem requires all these distributions to converge to a Gaussian. Depending on the nature of the tails, Berry-Eseen theorem dictates the upper limit to the number of iterations before these distributions attain their Gaussian limits [211, 212]. The approach to long time asymptotic diffusion limit can be described with Eq. 2.33. The existence of a disorder-specific long-time tail has been studied extensively by Haus and Kehr [213] and could tell us more on the nature of obstructions in the WM. Recently there has also been some studies on fractal nature of diffusion in brain tissue [214]. Considering tissue micro-structures as solid scaffold, time dependent diffusion coefficient throws light on the time required to explore all heterogeneous structures.

In contrast to high b-factor measurements, low b-value DWI requires short echo times and allows high SNR and better parameter estimates. In addition to obtaining tensor estimates at long $\Delta$, when DTI is measured at more than six directions it also throws light on the goodness of fit to the tensor model. The tensor model assumes Gaussian diffusion in all directions which is the low-spatial frequency approximation to the true PDF and inherently assumes that the three dimensional diffusion propagator to be a multivariate Gaussian. For the three dimensional propagator to be a multivariate Gaussian, not only each individual propagator at every direction needs to be a Gaussian but it is also necessary that the isodensity contour line - the locus of points in the 3-dimensional space with equal probability, should be an ellipsoid (see Fig. 2.1). It is this ellipsoid that is related to the diffusion tensor measured. The DTI protocol containing many diffusion directions can be used to estimate the goodness of a tensor fit. A good tensor fit suggests that the three dimensional diffusion propagator is close to a multivariate Gaussian. We define the non-Gaussianity (W) of the tensor model as ratio of the squared difference between the observed signal $S_o$ and the estimated signal $S_e$ from the tensor fit and the standard deviation of the noise $\sigma$.

$$W = \frac{(S_o - S_e)^2}{\sigma^2} \tag{6.7}$$

In a heterogeneous medium the move towards Gaussianity with increasing time is closely linked to the exchange between heterogeneous compartments. If the tensor fit improves with increasing $\Delta$, it suggests mixing between heterogeneous compartments. Similarly, worsening of the tensor fit signifies that at this time scale the restrictive barriers still influence the diffusion propagator and exchange is not yet relevant at the time range probed. This $\sigma$ was estimated from a region in the background where no signal is present and the noise is distributed with Rayleigh distribution. Since the signal at all data points were more than 3 times the noise threshold, the noise could be considered as normally

distributed.

### 6.2.2 Method

Two volunteers were scanned for this study. The diffusion times ($\Delta$) used were 50, 75, 150, 300, 675, 1500 and 2000 ms. The echo times for long $\Delta$ measurements were shortened to gain SNR. For all the measurement times the echo time varied between 33 and 48 ms. To accommodate many slices for long $\Delta$ measurements, the TR had to be varied between 4.5 and 12.6 seconds. Since $T_1$ for WM at 7 T is around 1500 ms, difference in TR times should not affect the results. Even with relatively long TR we could only acquire 6 slices for the longest diffusion time of 2 s. Six orthonormal diffusion weightings were applied at b-value of 0.15 ms/$\mu$m$^2$ and 60 diffusion weightings were applied at a b-value of 1.2 ms/$\mu$m$^2$. The six gradients used for 0.15 ms/$\mu$m$^2$ were applied at [110], [1–10], [101], [–101], [011], [01–1] directions, so that direct measurement along the z-axis could be avoided. The 60 directions were uniformly distributed over a hemisphere based on electrostatic repulsion [215].

Diffusion tensor estimates and subsequent data analyses were all performed using home built tools in Matlab(R) (Mathworks, Natick, MA). Only voxels with an SNR $\geq$ 10 were considered.

### 6.2.3 Results

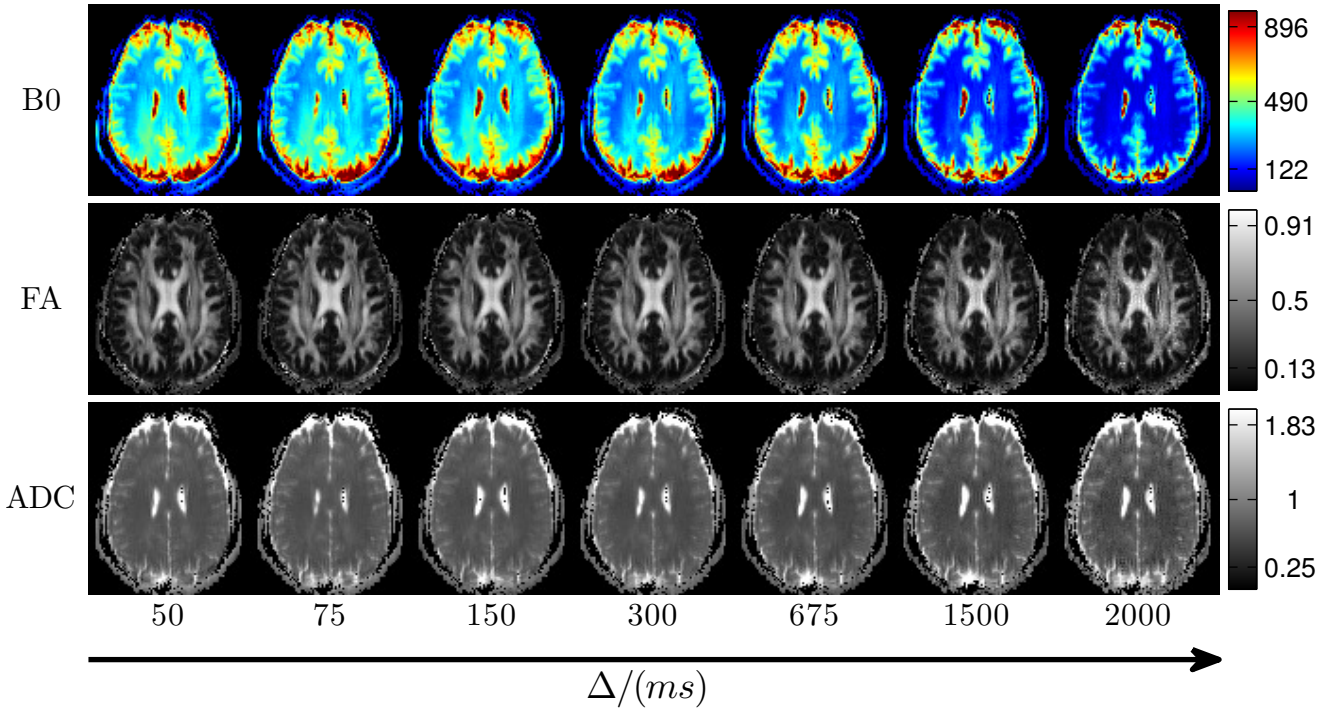



FIGURE 6.2: Maps of parameters obtained from tensor fit at different diffusion times; (i) non diffusion weighted signal ($B_0$, top row), (ii) Fractional anisotropy (middle row) and (iii) ADC (bottom row).

| | Subject 1 | | | | Subject2 | | | |
|---|---|---|---|---|---|---|---|---|
| $\Delta$ | ADC | $\lambda_1$ | $\lambda_2$ | $\lambda_3$ | ADC | $\lambda_1$ | $\lambda_2$ | $\lambda_3$ |
| 50 | $68 \pm 4$ | $132 \pm 16$ | $42 \pm 8$ | $31 \pm 7$ | $69 \pm 10$ | $137 \pm 24$ | $40 \pm 10$ | $29 \pm 9$ |
| 75 | $67 \pm 4$ | $129 \pm 14$ | $42 \pm 8$ | $31 \pm 7$ | $68 \pm 9$ | $136 \pm 24$ | $39 \pm 10$ | $28 \pm 9$ |
| 150 | $68 \pm 6$ | $133 \pm 17$ | $41 \pm 9$ | $29 \pm 8$ | $66 \pm 8$ | $132 \pm 22$ | $38 \pm 9$ | $27 \pm 7$ |
| 300 | $64 \pm 4$ | $127 \pm 14$ | $39 \pm 8$ | $28 \pm 6$ | $63 \pm 6$ | $127 \pm 20$ | $36 \pm 8$ | $26 \pm 7$ |
| 675 | $62 \pm 4$ | $121 \pm 14$ | $38 \pm 8$ | $27 \pm 7$ | $60 \pm 6$ | $120 \pm 18$ | $35 \pm 8$ | $24 \pm 6$ |
| 1500 | $61 \pm 7$ | $120 \pm 16$ | $38 \pm 9$ | $26 \pm 8$ | $59 \pm 8$ | $116 \pm 18$ | $36 \pm 8$ | $24 \pm 8$ |
| 2000 | $60 \pm 7$ | $118 \pm 16$ | $37 \pm 9$ | $25 \pm 8$ | $59 \pm 9$ | $116 \pm 20$ | $37 \pm 10$ | $25 \pm 9$ |

TABLE 6.1: Mean and standard deviation of the three eigenvalues and ADC. The values are in units of $10^{-2}$ $\mu$m$^2$/ms

The noise in background region was consistent across all diffusion times and for both the subjects. It had a mean of around 8.6 and a standard deviation of 4.6. With zero background signal, the Rayleigh noise in this region is around 7. Fig. 6.2 shows maps of non-diffusion weighted signal, $B_0$, fractional anisotropy FA, and ADC obtained for a slice passing through the mid corpus callosum. With increasing $\Delta$, a drop in signal is evident in $B_0$ (Fig. 6.2, top row), but FA and ADC values remain fairly consistent even though the longest $\Delta$ of 2000 ms is 40 times longer than the smallest $\Delta$ of 50 ms. Voxels with coherent white matter bundles show high FA, voxels that contain multiple WM fibers in different directions show a decreased FA, and FA in GM are even lower. The ADC values are quite similar between GM and WM. The plot on the left side of

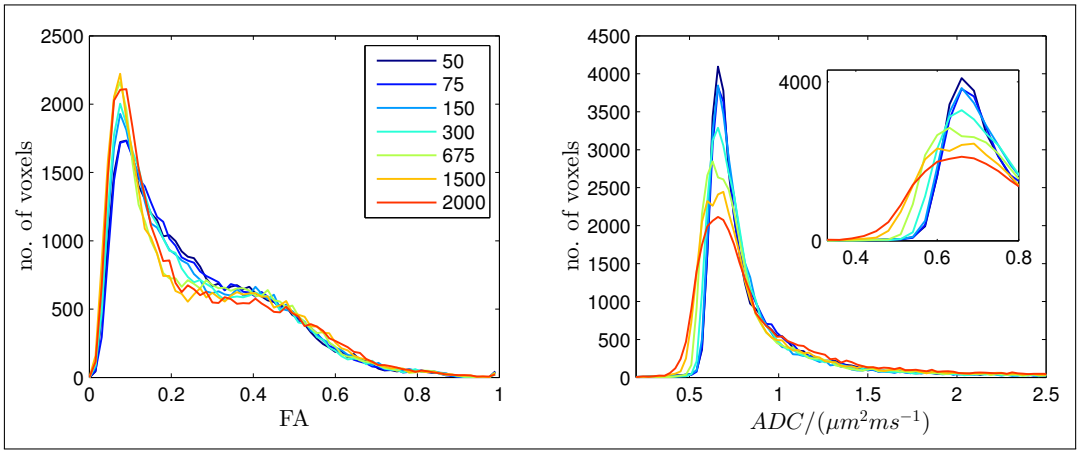



FIGURE 6.3: Normalized histograms of FA and ADC at different diffusion times.

Fig. 6.3 shows histograms for the FA for all the voxels in the brain. We see that FA is relatively stable across all diffusion times. The ADC shows a small but noticeable decrease with increasing diffusion times. Few voxels have the trace of the ADC greater that 1 $\mu$m$^2$/ms - these voxels are ones which have CSF partial volume. These voxels were excluded from further analysis. Only voxels that were selected as containing single WM bundle are reported in the findings.

| $\lambda_1$ | | | | | | | |
|---|---|---|---|---|---|---|---|
| $\Delta$ | 50 | 75 | 150 | 300 | 675 | 1500 | 2000 |
| 50 | – | .97 | .96 | .94 | .91 | .80 | .76 |
| 75 | .83 | – | .96 | .95 | .92 | .82 | .76 |
| 150 | .88 | .74 | – | .96 | .94 | .85 | .79 |
| 300 | .94 | .81 | .92 | – | .96 | .88 | .81 |
| 675 | .86 | .75 | .94 | .90 | – | .89 | .84 |
| 1500 | .80 | .68 | .89 | .85 | .90 | – | .82 |
| 2000 | .77 | .71 | .75 | .80 | .76 | .74 | – |

| $\lambda_2$ | | | | | | | |
|---|---|---|---|---|---|---|---|
| $\Delta$ | 50 | 75 | 150 | 300 | 675 | 1500 | 2000 |
| 50 | – | .86 | .74 | .64 | .57 | .30 | .29 |
| 75 | .67 | – | .81 | .75 | .69 | .49 | .36 |
| 150 | .81 | .51 | – | .86 | .80 | .49 | .40 |
| 300 | .90 | .69 | .83 | – | .87 | .55 | .45 |
| 675 | .77 | .59 | .86 | .84 | – | .61 | .53 |
| 1500 | .65 | .39 | .80 | .70 | .82 | – | .53 |
| 2000 | .59 | .47 | .59 | .62 | .59 | .62 | – |

| $\lambda_3$ | | | | | | | |
|---|---|---|---|---|---|---|---|
| $\Delta$ | 50 | 75 | 150 | 300 | 675 | 1500 | 2000 |
| 50 | – | .84 | .83 | .67 | .61 | .36 | .30 |
| 75 | .66 | – | .87 | .80 | .68 | .49 | .32 |
| 150 | .77 | .41 | – | .81 | .76 | .54 | .37 |
| 300 | .89 | .68 | .79 | – | .80 | .57 | .41 |
| 675 | .76 | .53 | .83 | .82 | – | .65 | .51 |
| 1500 | .61 | .30 | .79 | .65 | .80 | – | .53 |
| 2000 | .55 | .42 | .54 | .59 | .54 | .55 | – |

| ADC | | | | | | | |
|---|---|---|---|---|---|---|---|
| $\Delta$ | 50 | 75 | 150 | 300 | 675 | 1500 | 2000 |
| 50 | – | .90 | .86 | .79 | .71 | .58 | .57 |
| 75 | .54 | – | .87 | .86 | .77 | .68 | .59 |
| 150 | .71 | .30 | – | .86 | .82 | .70 | .62 |
| 300 | .81 | .52 | .72 | – | .85 | .74 | .65 |
| 675 | .65 | .35 | .75 | .69 | – | .74 | .68 |
| 1500 | .60 | .27 | .77 | .63 | .81 | – | .62 |
| 2000 | .55 | .35 | .53 | .61 | .55 | .57 | – |

TABLE 6.2: Voxel to voxel Pearson's Correlation coefficient of the three eigenvalues and the ADC. In each table the lower triangle is from Subject 1 and the upper triangle is from Subject 2. High degree of correlation between values at different diffusion times suggest that most of the variation is intrinsic to the voxel and not due to noise in parameter estimation.

**WM voxels** Table 6.1 shows mean ADC and the three eigenvalues for all diffusion times for the selected voxels. There is little difference between the tables obtained for the two volunteers. With increasing $\Delta$ a small drop in all three eigenvalues is observed. There seems to be a bigger drop in $\lambda_1$ between 150 and 675 ms. The decrease in $\lambda_2$ and $\lambda_3$ is small and no drastic changes could be observed. Mean value of $\lambda_1$ dropped from 1.35 $\mu$m$^2$/ms at 50 ms to about 1.17 $\mu$m$^2$/ms at 2 s. Similarly $\lambda_2$ decreased from 0.41 $\mu$m$^2$/ms to 0.37 $\mu$m$^2$/ms and $\lambda_3$ also decreased by about 0.05 $\mu$m$^2$/ms.

The results suggest that the variation between voxels is mostly due to inherent biological heterogeneity and not due to noise in the data. Table 6.2 shows the Pearson correlation coefficient between different $\Delta$s for the three eigenvalues. For a given $\Delta$, most of the variation observed in eigenvalues could be accounted by biological heterogeneity between the selected voxels. For example, after removing mean value from each voxel, the standard deviation of $\lambda_1$ at $\Delta = 50\,ms$ for the first volunteer was reduced from 0.16 to 0.03.

With time there is a significant decrease in all three of the eigenvalues. However, the changes are very small and the long time tail is very close to being a Gaussian for all eigenvalues. Fig. 6.4 shows the plot of the major eigenvalue ($\lambda_1$) against diffusion time. A power law fit shows that the best fit is an exponent of -0.04 which is very close to zero. Similar fits for the second and the third eigenvalues also had exponents equal to -0.03

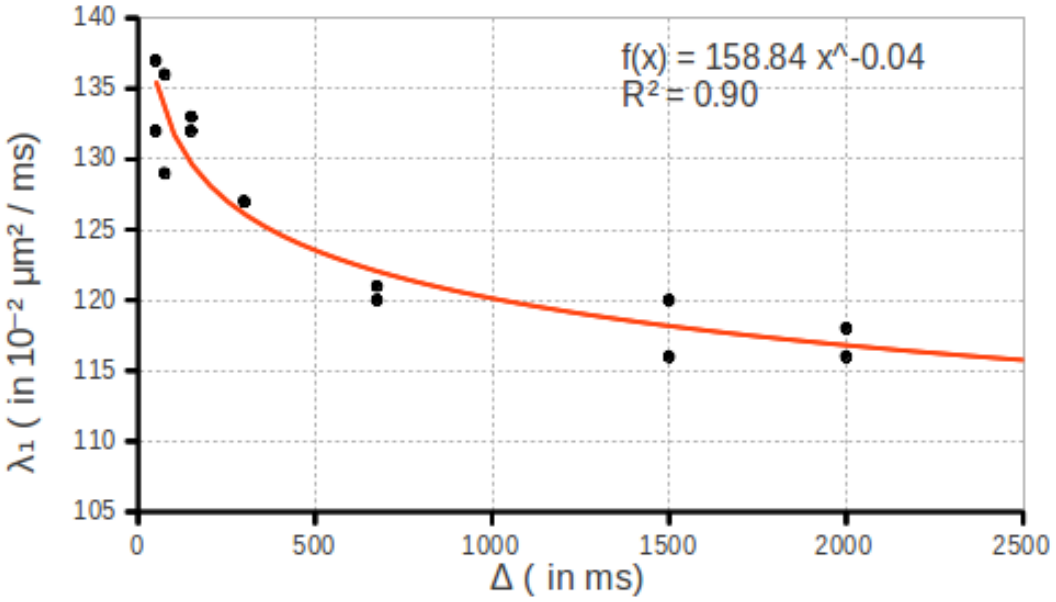



FIGURE 6.4: Plot of the first eigenvalue against diffusion time. There is some time dependence but the dependence is rather weak and does not allow for any conclusions to be drawn about the disorder in the system.

and -0.05. The time dependence is rather weak and does not allow for any conclusions to be drawn about the disorder in the system.

### Move towards Gaussianity

Fig. 6.5 shows the non-Gaussianty index for the different $\Delta$s. It is simply the misfit to the tensor model. The index assumes that the Rician noise threshold in a voxel does not change between measurements. Since a total of 66 measurements were used to estimate the tensor, a good tensor fit would be expected to have a value of around 66. The tensor fit gets better with increasing $\Delta$. This suggests that with time the three dimensional diffusion propagator becomes more and more Gaussian-like.

## 6.2.4 Discussion

To our knowledge this is the first human DTI study at very long diffusion times. The findings are similar with recent in vivo studies by Rane et al. in monkey brain in vivo [209]. Although their longest diffusion time was only 200 ms, they observe slight decrease in all three eigenvalues with increasing diffusion time. A similar study by Nair et al. in mice brain reported a decrease in $\lambda_3$ for diffusion times between diffusion time of 30 and 280 ms but stated that they could not see significant difference in $\lambda_1$.

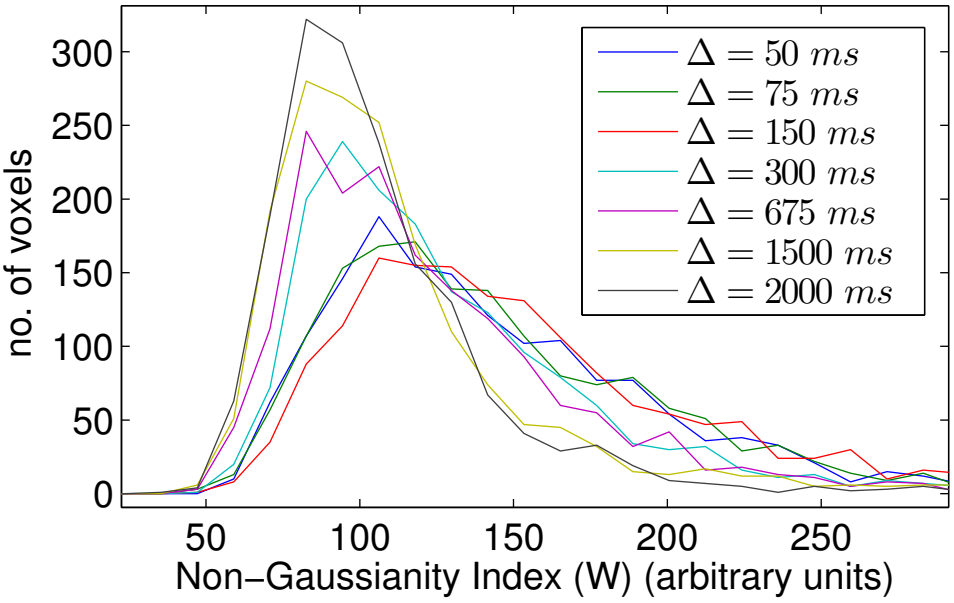



FIGURE 6.5: Histogram of non-Gaussianity index at differnt diffusion times.

An early study by Horsfield et al. [204] reported a decrease in ADC for all three directions in the corpus callosum. Their measurements were only done for three primary axes (x, y, z) assuming a known fiber direction, x, in the corpus callosum. While Horsfield et al. measured $\Delta$s between 40 and 800 ms, it was a constant gradient setup. Thus, the b-values used differed between different diffusion time. With $\delta = 20$ ms and a maximum gradient strength of 10 mT/m, the smallest $\Delta$ of 40 ms results in maximum possible b-factor around 0.053 ms/$\mu$m$^2$. Although intravoxel incoherent motion due to perfusion is minimal in white matter, at such small b-factors it could be the contributor for such an increased ADC. Indeed their results show very high ADC values for small $\Delta$s; at $\Delta = 40\,ms$ the ADC perpendicular to the corpus callosum has been reported to be more that 1 $\mu$m$^2$/ms. At their longest $\Delta$ at 800 ms, the maximum b-factor was around 1.25 ms/$\mu$m$^2$ which is comparable to the maximum that we used for our study. Ignoring results at short $\Delta$s, our findings are consistent to the findings by Horsfield et al. [204].

A study by Moonen et al. in cat brain in vivo is the only other study known to us that studied diffusion at $\Delta = 2000$ ms [113]. Their measurement was limited to only one direction and did not clearly distinguish between GM and WM. Nevertheless, it was reported that the diffusion coefficient did not vary greatly with time. We also observe that a drop in ADC is only about 0.1 $\mu$m$^2$/ms. Considering improvements in the scanner hardware and SNR over two decades, it is quite conceivable that their measurements were not sensitive to these small changes. Their measurements also included regions from both gray and white matter that could have further jumbled the data.

Since differences between the parameters are very small, a possible explanation could

also be sought in the way noise affects tensor estimation [216, 217]. Using Monte Carlo simulation and perturbation theory, Anderson had shown that noise in the signal could easily underestimate the smallest eigenvalue and overestimate the biggest eigenvalue [216]. During diagonalization of the tensor, decreased SNR leads to a systematic overestimation of tissue anisotropy [217]. However, this seems to not be the case in these measurements. The SNR was well above the noise level and we also find a decrease in all the three eigenvalues with time.

A small steady decrease in the diffusion coefficients with increasing $\Delta$ suggests that the presence of membranes and other structures act as weak hindrances for diffusive motion of protons. This is reminiscent of diffusion in polymers where non-Gaussian displacement PDFs are present due a broad range of fluctuations in the medium [210, 218]. Since the diffusion length for water molecules at 30 ms largely exceeds the average length scale of the micro-structures in the brain WM, time dependence of the ADC at long times can be qualitatively explained by assuming WM to have long range correlations, or fractal structures [214]. Recently, there has been much discussion about fractal nature of brain tissue [214]. Similar ADC time dependence in GM has recently been attributed to its underlying fractal structure [219]. However, especially in the direction radial to the axons, there is no such characteristic length scale that can be used to describe diffusion of water in brain WM. The dependence of ADC with time seems to be rather weak and the power exponent is also very small (Fig. 6.4). This raises a question on using such description to understand the present data.

A drop in diffusion coefficient could indicate presence of some long range correlations in the tissue that is reflected in diffusion data. But, such long range correlations in the spatial distribution of axonal diameters or pair correlations between axons have not been reported in the literature so far. If there actually exists some long range correlations, then wider range of diffusion times would perhaps allow us to further understand this phenomenon. Unfortunately DWI at longer diffusion times would hit SNR barrier due to tissue $T_1$ relaxation.

We also find that the non-Gaussianity index got smaller with increasing $\Delta$. The non-Gausssianity index indicates that the three dimensional propagator is slowly becoming more and more Gaussian-like. This means the heterogeneous compartments have already started mixing within the diffusion times observed in this study.

### 6.2.5 Conclusion

In this study, diffusion in human brain WM in vivo was measured at long diffusion times. It was found that with increasing $\Delta$, there is some decrease in the diffusion coefficients in all directions. Although these differences are significant at group level, they were not big enough to justify any treating WM diffusion from a fractal perspective.

For this data, one can conclude that the diffusion characteristic of the brain WM is relatively consistent over all diffusion times. The results show that if non-monoexponential diffusion persists at these long diffusion times, it probably is not just due to random disorder in the media. Another important aspect of this measurement is decreasing non-Gaussianity index with increasing $\Delta$. This is a qualitative indication that although the exchange time is long, it is within the reach of measurements at these long diffusion times. This signifies that measurements at long diffusion times should take exchange into consideration.

The following sections will first assess the validity of different mathematical models in fitting diffusion decay curves at high b-values. After identifying a working model, optimization of the DWI protocol for high b-value measurements is described. The final section describes experiment performed at long diffusion time and high b-values.

## 6.3 Fitting diffusion decay curves *in vivo* at high b-factors

### 6.3.1 Introduction

High b-factor ( >2 ms/$\mu$m$^2$ ) DWI in brain tissue is known to have a non-monoexponential decay . Non-monoexponential diffusional decay can be observed in both gray and white matter regions but given the gross difference in the tissue architecture, the origin of such behavior might be very different. DWI at high b-factors is sensitive to tissue properties that are not revealed at lower b-values. ADC, for instance, is quite similar in both GM and WM, however at high b-factors the signal decay is faster in GM than in WM [121, 220]. Due to its sensitivity towards smaller length scales, high b-factor DWI promises to reveal properties of tissue substructures that are histologically relevant. Compared to GM, the micro-structural geometry of WM is quite simple. This is especially true for WM comprised of a single fiber bundle. In principle, diffusional decay in WM should be simpler to model, but a complete consensus picture is yet to be formed.

The straight-forward approach to characterize diffusion is to look at the diffusion propagator itself [221]. In an ideal PFG-NMR experiment with narrow diffusion gradients, the signal intensity of a PFG-NMR experiment can be related to the diffusion propagator through the Fourier relation

$$S(q,\Delta) = S_0 \int P(R,\tau) e^{iqR\ dR} = S_0 F[P(\Delta r)] \tag{6.8}$$

where $S_0 = S(0,\Delta)$ is the signal without any diffusion weighting, and $R$ is the spin displacement, $\Delta$ is the diffusion time and $q = \gamma\delta G/2\pi$ is the gradient integral [34]. The ability to resolve microscopic lengths is reminiscent of an imaging experiment and referred as q-space imaging. For finite gradients used for *in vivo* imaging the echo attenuation should be considered using Mitra and Halperin's center of moment description [104].

Many high b-factor DW studies have explored q-space imaging based techniques. Early adapters of this technique were Assaf and Cohen [140] who extensively used q-space analysis to throw light on non-Gaussian diffusion in WM. It was expected that a q-space based measurement perpendicular to the WM fiber bundle would reveal information about the diameters of the axons. However, they were unable to realize diffusion diffraction patterns in brain WM and could not unequivocally establish a case for restriction in brain WM *in vivo* [122]. It was soon understood that even if each axon does restrict water, a voxel comprised of a broad distribution of axons could not produce diffusion-diffraction pattern [177]. A three dimensional version of q-space imaging in brain tissue has been widely referred to as diffusion spectrum imaging (DSI) [222]. DSI is the three-dimensional Fourier transform of the diffusional decay curves. It considers contours in the propagator as arising from different fiber populations and tries to resolve multiple

fiber orientations within a voxel. DSI has been claimed to reveal many histologically relevant information about gross WM organization, but it does not probe into properties of the propagator and its relation to tissue architecture. Although q-space imaging does contain information about tissue micro-structure at microscopic scale, it only provides a blurred picture of the composite tissue substructures. Thus, it is necessary to use model that is simple and yet captures essential features responsible for the observed signal decay.

Most diffusion studies in brain WM model signal decay as arising from multiple Gaussian components. In a Gaussian mixture model the tissue is supposed to be composed of multiple Gaussian compartments so that the resulting DW signal against gradient squared (b-factor) is described as a linear sum of multiple exponential decays. The simplest trivial case for the Gaussian model is the monoexponential decay which assumes a homogeneous spin population. A biexponential fit is the simplest fit that describes the Gaussian mixture model. It assumes two water pools where the propagator for each pool is a Gaussian. Many studies that employ high b-factor DWI *in vivo* have reported a biexponential fit as a chosen analysis method. A biexponential fit is easy to employ, requires less data points than a Fourier transform and fits most available data surprisingly well [119, 121]. It estimates four parameters, the non-diffusion weighted signal $S_0$, diffusion coefficient of the fast water pool $D_f$, diffusion coefficient of the slow water pool ($D_s$) and the relative volume fractions of the slow and fast pool.

When the ADC of the slower pool ($D_s$) is close to zero ( $b_{max} * D_s < 0.1$ ), the two component system could be effectively described with one less parameter ($D_s = 0$). Such a fit is especially relevant to diffusion measurements perpendicular to the fiber direction if intra-axonal water is completely restricted by the myelin walls. In this case, the diffusion coefficient for the intra-axonal water would be close to zero. We call this fit MwB - monoexponential with a baseline.

A small water fraction with a very small diffusion coefficient ($D \sim 10^{-2}$ $\mu m^2/ms$) has also been reported in literature [76, 115, 116]. This ultra slow water fraction has only been reported in studies employing very high b-factors ($>50$ $ms/\mu m^2$). For moderately hight b-values attainable for human *in vivo* measurements this very slow fraction of water is equivalent to still water fraction that does not seem to be moving. The volume fraction reported for this water pool is also rather low ( $< 10\%$). At low SNR, it could be easily mistaken for Rician noise. If this fraction is indeed present, it will cause the signal to attain a baseline level in addition to the biexponential fit. To accommodate for this situation, we can extend the biexponential fit to include a third water pool with zero diffusion coefficient. This fit is referred as BwB fit - biexponential with baseline. The BwB fit requires one more parameter than the biexponential fit.

Recently, there have been efforts to describe diffusion data in brain WM with fits that do not describe the data as arising from deviations in Gaussianity of the diffusion

propagator. The two ways to do so are kurtosis fit [44] and the stretched exponential fit [214, 223]. Both fits are more parsimonious than the biexponential fit and need one less parameter. There have been reports in literature claiming these fits to be as effective as a biexponential fit [45, 223].

The kurtosis fit is the second order cumulant expansion of the log of the diffusion decay signal and measures deviation of the diffusion propagator from a Gaussian. The kurtosis fits assumes that diffusion processes is inherently Gaussian with small deviations due to the medium. These small deviations could be explained with the kurtosis of the diffusion propagator. It does so by fitting a second order term to the logarithm of the signal decay against b-value.

$$ln(S) = ln(S_0) - bD - k(bD)^2 + ... \tag{6.9}$$

where $S$ is the signal, $S_0$ is the unweighted signal and $k$ is the kurtosis term. The stretched exponential fit describes the signal decay with a power of $bD$.

$$S = S_0\, exp(-(bD)^{\alpha}) \tag{6.10}$$

where $S$ is the signal, $S_0$ is the unweighted signal and $\alpha$ is the stretched exponent. When $\alpha$ is unity this fit reduces to a monoexponential fit.

### 6.3.2 Rician noise correction

At high b-factors and adequately high spatial resolution, the SNR is quite low and repeated measurements are required. Magnitude MR images follow Rician noise distribution which causes an upwardly biased mean for low SNR data [224]. If not accounted for, this upward bias increasingly affects high b-factor images and distorts diffusional decay curves [225]. An unbiased estimate of the 'true' signal from repeated measurements is given by the maximum likelihood estimate (MLE) of the probability density function. MLE in the context of MRI has been investigated by Sijbers et al. [226]. In the presence of phase errors in MR images, MLE from magnitude data is significantly more reliable than from complex data [227].

For the Rician distribution, given the true underlying signal $A$ and the standard deviation of noise $\sigma$, the probability of observing $S$ is defined by the probability density function

$$p(S|A,\sigma) = \frac{S}{\sigma^2} e^{\frac{S^2+A^2}{\sigma^2}}\, I_0\left(\frac{AS}{\sigma^2}\right) u\,(S) \tag{6.11}$$

where $I_0$ is the modified zeroth order Bessel function of the first kind, and $u$ is the Heaviside function ensuring non-negativity of $S$. The log-likelihood, $L$, of the observed

data, $A$, can be calculated as

$$log\ L(\hat{A},\hat{\sigma}) \sim \sum_{i=1}^{N} logI_0\left(\frac{AS}{\sigma^2}\right) - \sum_{i=1}^{N}\frac{A^2}{2\sigma^2} \tag{6.12}$$

$L$ attains its global maximum at the MLE of the data. Once MLE has been calculated for each b-factor, further fitting with the appropriate model can be performed.

A good fit to the data is precursor to a proper model description. The first goal of this study is to compare the goodness of different fitting procedures applied to diffusion decay data obtained from measurements along a single gradient direction. Fittings can also be performed for diffusion decay data obtained in all directions but single direction fits are less messy and allow easier interpretation data and the fitting parameters. We report on the high b-value DWI for three orthonormal directions (x,y,z) in the scanner coordinate system. MwB, biexponential, BwB, kurtosis and stretched exponential fits were performed to the same data and their effectiveness were compared. After evaluating the fitting results, more analyses was done for the suitable fitting procedures.

### 6.3.3 Methods

Four volunteers were scanned for this study. The parameters used were: $TR/TE/\Delta/\delta =$ 3000/59/121/16 ms. Images were acquired with 2 mm isotropic resolution using 2D single shot DW-STEAM-EPI. The acquisition bandwidth was 1930 Hz/pixel and the k-space was partially sampled at 5/8.

The acquired region of interest extended superiorly from the dorsal surface of the corpus callosum. 16 diffusion weightings were applied in the anterior-posterior direction for all three gradient directions. Gradients were incremented linearly in 16 steps from 6% to 95% of the maximal gradient strength such that b-values increased quadratically starting from 0.03, 0.14, 0.27, ..., 7.07, and 8.0 ms/$\mu$m$^2$. For each b-value the measurement was repeated 16 times. A total of 256 measurements for each diffusion direction and 768 total scans took about 39 minutes.

Additionally a DTI protocol was also acquired with a shell of 60 directions at b-value of 1.2 ms/$\mu$m$^2$ and 6 directions at 0.15 ms/$\mu$m$^2$. All other parameters for this DTI protocol were exactly same as the high b-factor protocol. The DTI measurement was used to select WM voxels containing single WM fiber bundles. The selection procedure was same as described in the previously [Section 6.1.4]. After FA threshold. only voxels present were those in the corpus callosum, cerebrospinal tract and cingulum that are the three major white matter fiber bundles.

All subsequent analyses with the images were performed using home-built tools in Matlab(R) (Mathworks, Natick, MA). First, for all b-values the 16 repeated measurements were used to compute the maximum likelihood estimate (MLE) of the true signal.

Maximum likelihood estimate is asymptotically normal and its variance is proportional to the inverse of the Fisher information matrix. The MLEs were used as input to fit the data. The reciprocal of variances obtained from the MLE were used as the weights for the fits. All fits minimized the chi-squared value

$$_{min}\left(\chi^2 = \Sigma_i^n\left[\frac{(S_i - So_i)^2}{\sigma_i^2}\right]\right) \tag{6.13}$$

where, $S_i$ is the expected signal from the fitting method, $So_i$ is the MLE of the data at each b-value, $\sigma_i^2$ are the variances of the MLEs and $n$ is the number of data points used for fitting. The number of fit parameters for the MwB, biexponential, BwB, kurtosis and stretched exponential fits were respectively 3, 4, 5, 3, and 3. Since the fitting procedure is non-linear, the estimated parameters are not linearly independent which results in slightly underestimated degrees of freedom.

### 6.3.4 Results

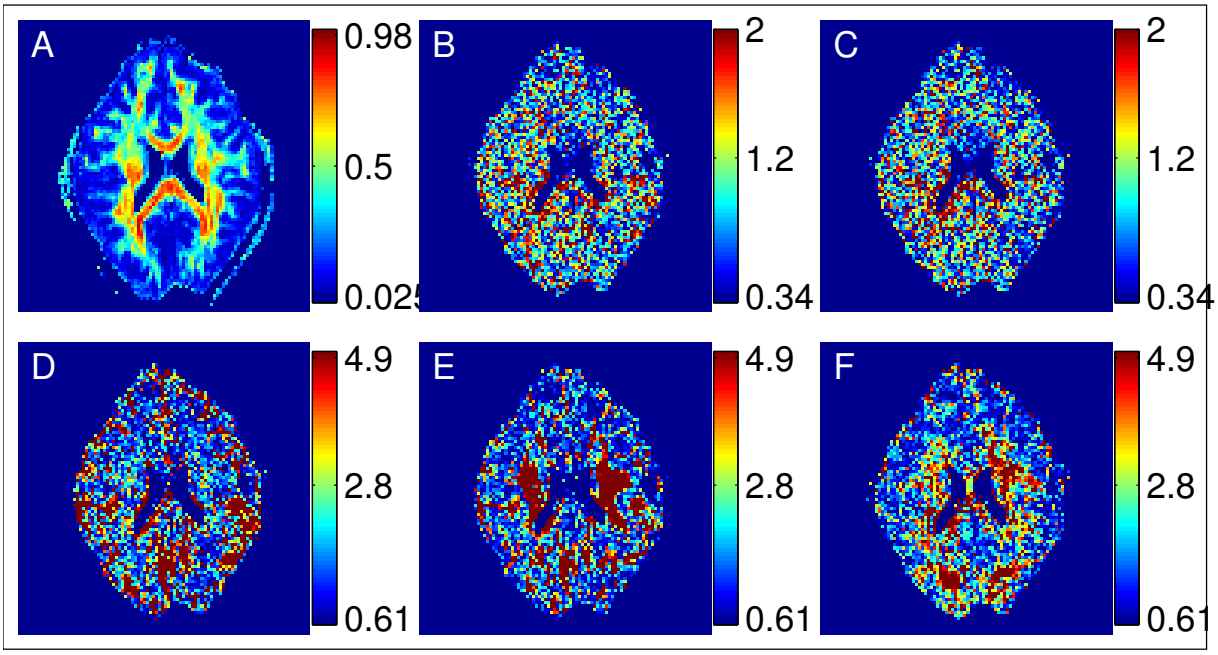



FIGURE 6.6: FA map obtained from DTI measurement (A) (shown for anatomical reference), and maps of reduced $\chi^2$ obtained from different fitting procedures: biexponential fit (B) , BwB fit (C) , MwB fit (D), kurtosis fit (E), and stretched exponential fit (F). Gradient were applied in the x-axis. The slice contains mid corpus callosum ( $\parallel$ to x-axis) and the cerebrospinal tract ( $\perp$ to x-axis). The biexponential and the BwB fits perform far better than the three other fits. The colorbars for the three lower maps had to be scaled up to show some contrast.

Fig. 6.6 shows maps of reduced $\chi^2$ obtained in a representative slice for different fitting procedures in a slice. When the fitting model sufficiently describes the data we expect

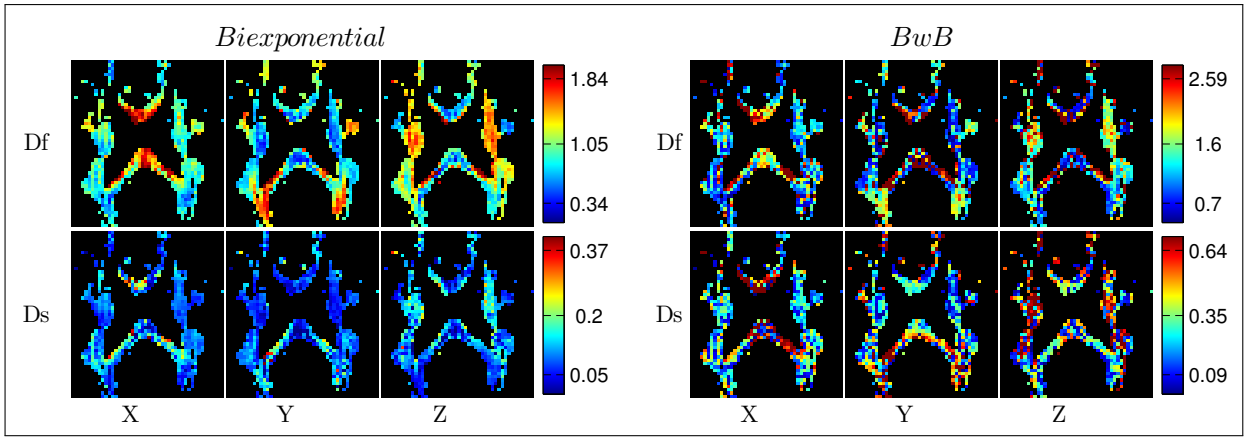



FIGURE 6.7: Diffusion coefficients of the fast component $D_f$ (top row) and the slow component $D_s$ (bottom row) components obtained from biexponential fit (left side) and BwB fit (right side) . The three columns are data from x, y, and z axes respectively. The units are in $\mu m^2/ms$. Data is only shown for voxels containing single WM bundle.

the reduced $\chi^2$ to be close to unity. When the values are below unity, the model could have over-fitted the data and some parameters could be superfluous.

The three parameter fits, BwB, kurtosis and the stretched exponential fits, do not fare well in explaining the data. For most voxels, both the kurtosis and the stretched exponential fit performed worse than a monoexponential fit with baseline.

With the addition of one extra parameter, the biexponential fit by far outperforms all three parameter fits. The reduced $\chi^2$ values are more than three folds smaller in all voxels. For most of the WM, the $\chi^2$ for the biexponential fits are close to unity suggesting that it is a good fitting method. When a baseline is added to the biexponential fit (BwB fit) the resulting values are smaller $\chi^2$, however, the improvement is only around 0.1. From $\chi^2$ values alone it is quite difficult to judge if an additional parameter is actually helpful.

The top row in Fig. 6.7 shows estimated $D_f$ for the biexponential fit (left side) and the BwB fit (right side). The three columns show parameters for x, y and z axes from left to right. Only voxels in the selected WM region are shown. The directional dependence of $D_f$ is seen for both maps: at the corpus callosum, for instance, diffusion coefficients are highest when measured along the x-axis.

For the biexponential fit the maps show excellent spatial homogeneity. Since all data processing is done at the voxel level, a smooth spatial maps of the parameters further adds credit to the effectiveness of the biexponential fit in describing the data. This is not true of the BwB fit: in a quarter of the voxels $D_f$ attains an unreasonably high value ($D_f > 3\ \mu m^2/ms$) Similarly in many other voxels $D_s$ is close to zero, signaling that $D_s$ and the baseline could be effectively combined together. Furthermore, in some voxels $D_f$ and $D_s$ were very close to each other. About 40% voxels were identified to

have one of the aforementioned issues. Subscribing to the parsimonious solution and taking spatial homogeneity into consideration, for this range of b-values a biexponential fit seems to be sufficient in describing diffusion decay curves in a single direction.

#### 6.3.4.1 Direction Dependence of Biexponential Parameters

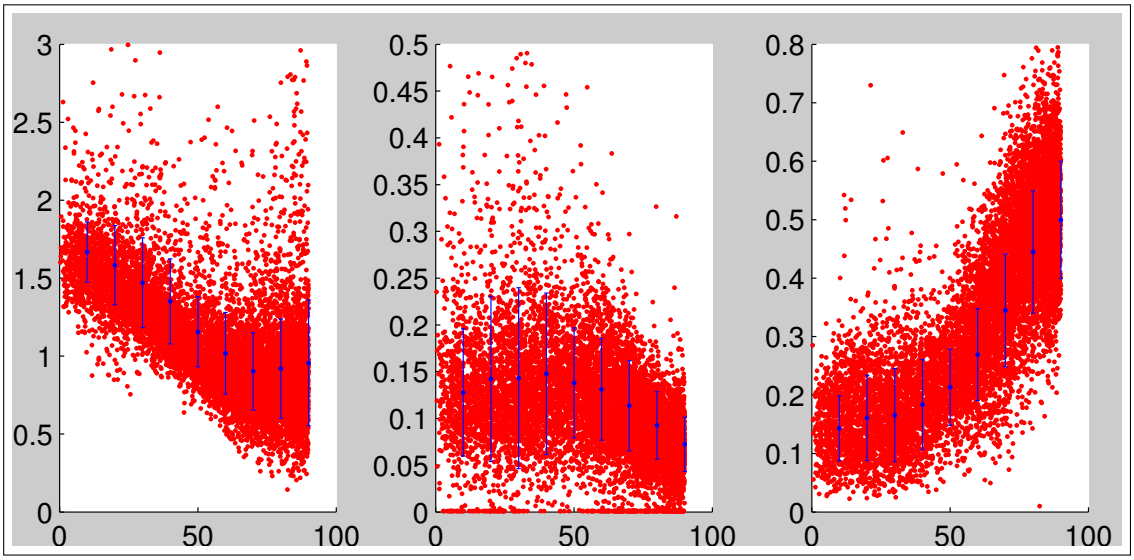



FIGURE 6.8: Dependence of biexponential fit parameters (from left to right: $D_f$, $D_s$ and $V_s$) on the angle between the fiber bundle and the direction of the diffusion measurement. The scatter plot are individual points from each voxel and the blue dot with error bar are mean values taken at 10° intervals.

To quantify the direction dependence of the estimated parameters angle we calculated the angle between the primary eigenvector obtained from the DTI measurement and the measured direction. The fit parameters show clear correlation with the primary eigenvector estimated from DTI measurement (Fig. 6.8). $D_f$ drops sharply from 1.65 $\mu m^2/ms$ when measured parallel to the fiber orientation to about 0.5 $\mu m^2/ms$ when measured in perpendicular direction. The slow diffusion coefficient $D_s$ is around 0.15 $\mu m^2/ms$ when measured axial to the fibers. It is relatively stable until about 50 degrees and then slowly drops with increasing angle.

Relative volume fractions of the two components also showed direction dependence. The volume fraction of the slow component in Fig 6.8(right) is about 0.15 in the parallel direction, the increase is relatively small till 50° and then rises rapidly reaching 0.55 in the perpendicular direction.

#### 6.3.4.2 Fitting 3 directions

The dependence of the relative volume fractions on the diffusion direction suggests that although biexponential fit satisfactorily describes the decay curves in one direction, it

would fail to fit diffusion decay curves in all directions. This encouraged us to check if BwB fit would be more relevant to fit all three direction. Since the three directions were orthogonal to each other $D_f$ and $D_s$ were allowed to vary between the three directions but with relative volume fractions kept the same. For comparison both BwB and biexponential fit were performed.

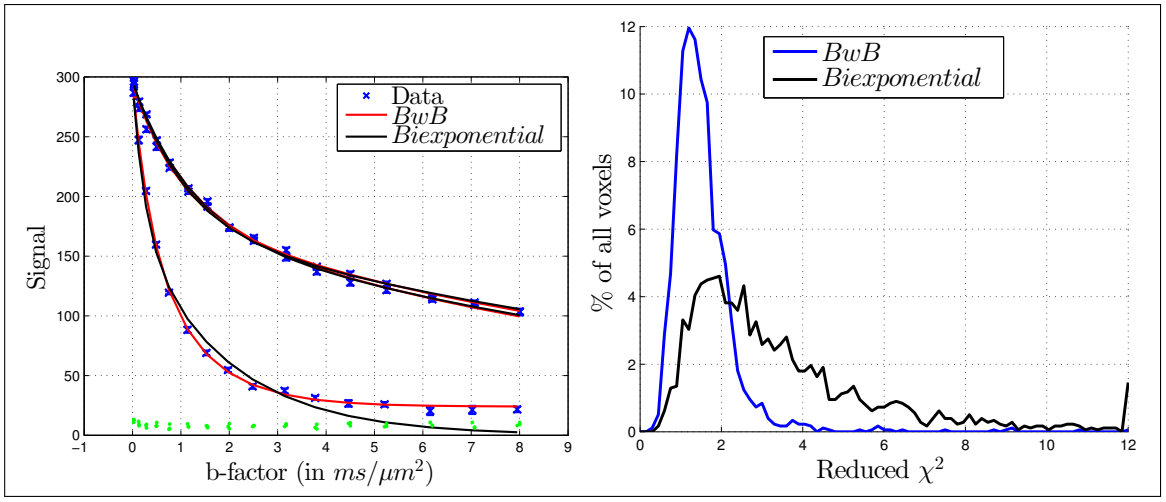



FIGURE 6.9: Left: Diffusion curve from a voxel in the cerebrospinal tract. The BwB is a good fit for all three directions, but biexponential fit fails to do so. The green dots are the Rician noise parameters. Right: Normalized histogram of reduced $\chi^2$ values from biexponential fit and BwB fit.

As expected, biexponential fit was no longer sufficient to describe the decay curves. The BwB performed much better than the biexponential fit and sufficiently fitted the data. To illustrate this, Fig. 6.9 (left) shows an example data taken from a voxel in the cerebrospinal tract. In this voxel, the biexponential fit which was excellent when fitted in individual directions now fails to describe the data in the parallel direction. BwB fit however is still relevant and is able to capture the essential features of the data. Fig. 6.9 (right) shows the histogram of reduced chi-squared values obtained for both of the these fits. For 87% of the voxels the reduced chi-squared value for the BwB fit was smaller than 2, whereas it was only 35% for the biexponential fit.

Fig. 6.10 (left) shows the $D_f$, $D_s$ obtained from the BwB fit for all selected WM voxels. Here both $D_f$ and $D_s$ show clear directionality. On the right, Fig. 6.10 shows the histograms of the relative volume fraction of the three water pools estimated from the BwB fit.

### 6.3.5 Discussion

The MwB fit, the kurtosis fit and the stretched exponential fit are clearly insufficient for the range of b-factors used in this study. Previous reports that compared biexponential fit to kurtosis fit [45] or to stretched exponential [223] fit obtained data for a limited

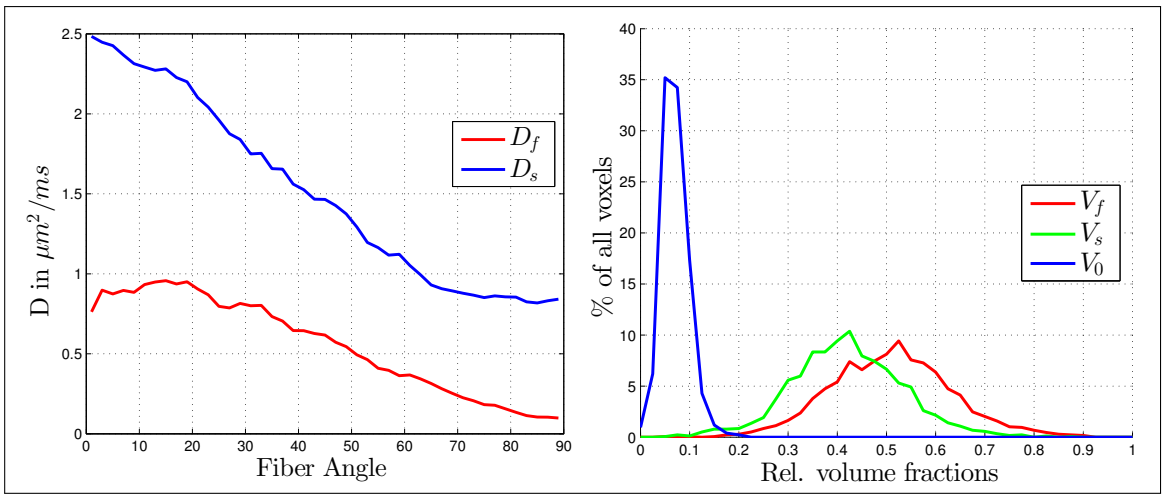



FIGURE 6.10: BwB fit parameters. Left: Dependence of $D_f$ and $D_s$ on the angle between the fiber bundle and the direction of the diffusion measurement. The horizontal axis shows angle made by the measurement direction with the primary eigenvector of the diffusion tensor. Right: Volume fraction of the fast component $V_f$, slow component $V_s$ and the baseline $V_0$.

range of b-values. In light of our results it seems that these fits could only be applied for very limited b-value range. Since these fits require just three parameters, they could be potentially used for a smaller range of b-factors but care should be taken to justify such fits to physical variables. More importantly, for the range of b-values used in this study, the biexponential fit is found to be a good way to describe diffusion decay curves in any single direction. Similar conclusions about the effectiveness of the biexponential fit have been argued in the literature before [121].

The biexponential fit was necessary even for measurements parallel to the fiber orientation, not even a MwB fit was enough to explain the data. To stress, if the non-monoexponential decay was simply due to noise threshold then the MwB fit would be enough to account for such baseline effects.

The biexponential fit could be heuristic, but it can be instrumental in elucidating deficiencies in different models. Recently there have been efforts to model axonal dispersion using diffusion data [228]. The necessity of a biexponential fit in parallel direction suggests that axonal orientation dispersion alone is **not** enough to describe such deviations from monoexponentiality. Leegaard et al. [229] recorded individual fiber orientations from myelin stained slices and estimated angular orientation distribution that resembled a Gaussian. In the corpus callosum the angular orientation distribution was reported to have a FWHM of 34°. Fiber dispersion allows some deviation from a monoexponential decay but it does not necessitate a biexponential fit. As shown in Fig. 6.11, allowing orientation dispersion to otherwise identical fibers effectively reduces the ADC by a mere 6% but the decay curve can still looks sufficiently monoexponential. It is seen

that the decay is fast and the signal is expected to go below 3% when b-factor is about 2.5 ms/$\mu$m$^2$. This is contradictory to our finding that there is still about 20% signal remaining at that b-value. Fiber dispersion is not the main reason of non-monoexponential behavior of the diffusion decay curve.

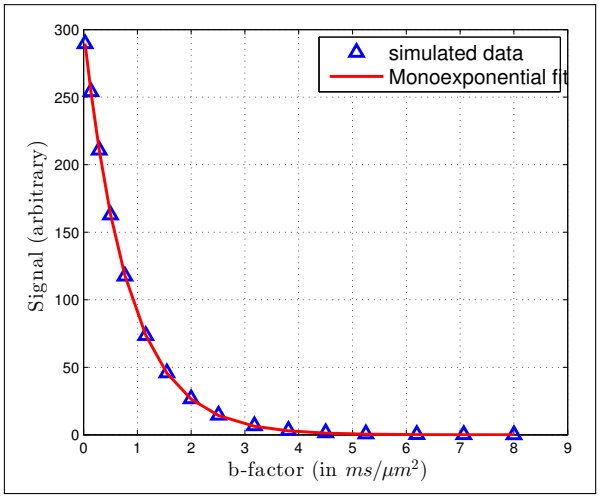



FIGURE 6.11: Simulated Data for a fiber dispersion (blue triangles). The diffusion coefficient parallel to the fiber is taken to be 1.3 $\mu$m$^2$/ms. Gaussian fiber dispersion with FWHM of 34° has been assumed. The red line is the exponential fit which fits very well to the simulated data.

It should also be noted that fiber dispersion might be responsible for a non-zero $D_s$ in the perpendicular direction. The reduced ADC in the parallel direction appears as increased ADC in perpendicular direction. A reduction from the parallel direction with 'true' fiber ADC of 1 would result in parallel ADC to be 0.94 which is a mere 6% reduction, but it would also increase the ADC perpendicular by the same amount. If the 'true' ADC perpendicular is close to zero then the measured ADC perpendicular would be close to 0.06 - entirely due to fiber dispersion.

The necessity of a biexponential fit to describe diffusion parallel to WM bundle is not new, but has been overlooked. Indeed, early reports like those from Assaf et al. [122] and Clark and Le Bihan [119] both reported biexponential decay for WM in parallel as well as in perpendicular directions. Similar to our findings, Assaf et al. reported that volume fraction of the slow component varied from 14% in the axial direction compared to around 40% in the radial direction [122]. However, this important finding was soon overlooked presumably due to an excitement to fully define WM diffusion in all directions. Another reason for this neglect was probably due to the feasibility of fitting a biexponential decay in many physically single compartments. Sehy et al. [230] showed that in Xenopus oocyte both fast and slow diffusion components arose from intracellular space. Similarly, using densely packed human erythrocytes, it was shown that a biexponential fit sufficiently described diffusion decay curves arising from a single

component system with hindrances [231]. The overarching goodness of fit is due to the exponential nature of decay curves.

Since the volume fractions differ between axial and radial slow components, the bi-exponential fit is only heuristic but not indicative of physical compartmentalization. Furthermore, neither of these volume fractions could be properly related to the tissue micro-structure [232]. The biexponential fit only works for single diffusion direction but clearly fails to describe diffusion decay curves for all three directions in a single go. The BwB fit performs much better in such a situation.

Applying BwB fit in 3 directions is actually very similar to measurements performed by Stanisz et al. [124] except for the exchange terms in their model (Section 4.3.3). Their measurements were also performed only in 2 directions parallel and perpendicular to the fiber orientation. These relatively simple measurements already show that since diffusion in perpendicular and parallel directions show very different characteristics, any Gaussian mixture model used to fit the data will require at least 3 compartments to get a good fit in all diffusion directions. It is therefore not surprising that recent studies performed in many directions with multiple b-values have also reported better performance with a 3 compartment model [148]. Mixture of exponential decays are indeed extremely difficult to distinguish and thus adding more parameters could cause data over-fitting as was observed for the BwB fit.

### 6.3.6 Conclusion

Does this mean that there are three compartments in the WM: two of them with relatively similar volume fractions, and one with a small volume fractions?

From this measurement, and the *ex vivo* measurements in the previous chapter, two similarities are seen for diffusion in the parallel direction. First, there is a clear deviation from the monoexponential decay. At short time myelin also contributed to the slow diffusion pool but even at longer echo times there still exist a small but significant slow water fraction. Second, there is no easy and meaningful way to attribute this non-monoexponential decay to a single compartment with some restricted boundaries, disorder in the medium or to the application of gradients with long duration. In *ex vivo* sample, the ILT showed that two clear diffusion compartments and in this measurement neither Kurtosis not stretched exponential fits could fit the data satisfactorily. Inclusion of $(bD)^3$ and other higher order terms could have fitted the data better, but the usefulness of those fits is difficult to justify.

For the perpendicular direction, myelin layers should restrict a radial mobility of a large water fraction residing inside the axons and in the periaxonal space. It seems that the slow water fraction found in the parallel direction is also slow in the perpendicular direction. This is reasonable from a physical point of view: at a diffusion time of 121 ms

employed in this study. The slow component in the parallel direction has a diffusion length of around 2.5 $\mu$m. If this pool was more mobile in perpendicular direction with greater diffusion length, it would mix with faster water pools.

The volume fraction of this slow pool is found to be around 18%. Excluding the contribution of myelin in the previous *ex vivo* study, the slow component has a slightly smaller but similar contribution. The two conjectures from the slow water pool (i) from the tubules and (ii) from the isotropically distributed glial processes, both satisfy the condition that the slow pool does not mix with other water.

Discrimination between these two pools is perhaps easier at longer times with high diffusion weighting. At high b-factors the signal from the water in glial processes would be isotropic and show a $1/\sqrt{b}$ dependence in all directions. At longer times, the diffusion length ($\sqrt{2Dt}$) of slow pool would exceed the length of most of the the tiny microscopic tubular structures. Mixing of the component should allow the decay curves to move towards becoming more monoexponential. Even, if this component somehow still survives, its diffusion decay should be monoexponential and should **not** be isotropic.

### 6.3.7 Moving Forward

At higher b-values the signal is further affected by the Rician noise threshold. This issue could be addressed with more repeated measurements or with bigger voxels. While repeated measurements cost extra time, bigger voxels are increasingly affected by partial voluming.

Along the diffusion decay curve, it would be desirable to obtain less repetitions for high SNR data and more for low SNR data. Additionally, performing MLE at each b-value requires many repetitions and is measurement time expensive.

In order to optimize DWI, two major issues are address in the following section (i) partial voluming and (ii) repeated measurements.

## 6.4 Improved DWI measurement and analysis for high b-factors

This section is devoted to estimating parameters from high b-factor DWI with acceptable scan time, tissue discrimination, accuracy and precision. There are two pertaining issues: (i) avoiding partial voluming and (ii) reducing number of repeated measurements along diffusion decay curves.

This chapter is devoted to two ways which can jointly be helpful in optimized DWI *in vivo* measurement.

### 6.4.1 b-filtering

Voxels lying in the border between two distinct tissues are prone to partial volume effects. Voxels lying near the cortical surface are especially prone to contamination from cerebrospinal fluid (CSF). Diffusion un-weighted images (bvalue = 0) are also prone to ringing artifacts near CSF boundaries and impact the accuracy of diffusion measurements. A common workaround has been to apply an additional RF pulse to invert the magnetization at a time TI before the excitation pulse [233]. With proper combination of TI and repetition time (TR), this technique, commonly known as fluid attenuation inversion recovery (FLAIR), allows suppression of spins with specific $T_1$ times. While CSF signal is suppressed, a significant fraction of white and gray matter is also attenuated. At 7 T, considering $T_1$ times to have values of 1.2 s, 2.0 s and 4.0 s for WM, GM and CSF respectively [234], complete CSF suppression would result in at least 20% signal loss in WM and 50% signal loss GM. This would further be reduced at shorter TR times. Moreover, at high fields, where constraints in the specific absorption rate (SAR) is often the limiting factor, using an extra inversion RF pulse requires a drastic increase in the total acquisition time.

An alternative way to crush CSF signal with little signal loss for WM and GM would be to apply diffusion weighting. At body temperature, CSF has a diffusion coefficient of about 3.0 $\mu m^2/ms$ and loses about 80% of its signal at a b-value of 0.5 $ms/\mu m^2$, while signal from brain tissue with a mean diffusion coefficient of about 1.0 $\mu m^2/ms$ is attenuated by only 40% at that b-value. Signal suppression based on b-value is henceforth described as b-filtering method. In DWI, a possible drawback for this method is its impact on ADCs that would be estimated from curves only at higher b-factors. However, if enough SNR is available, diffusion decay curves can be considered as sum of exponential terms and the ADCs obtained should not deviate far from their true values.

### 6.4.2 joint Maximum Likelihood Estimate

MLE is a good way to estimate the true underlying signal, but still underestimates the underlying signal for SNR below 3 [226]. At high b-values, low SNR would result in biased estimation of the true signal. However, this bias in ML estimator is mitigated when the noise standard deviation $\sigma$ is known a priori. If $\sigma$ can be considered independent of the diffusion weighting then the 'true' signal $A$, at different b-values can be expressed as a function of the diffusion weighting, the diffusion coefficients and their volume fractions, $A(b, \vec{D}, \vec{F}) = A_0 \sum_{j=1}^{2} F_j \ e^{-bD_j}$. In this case, the log-likelihood of the observed decay curve is

$$log\ L(\hat{A}_0, \hat{\vec{D}}, \hat{\vec{F}}, \hat{\sigma}) \sim \sum_{i=1}^{N} logI_0 \left( \frac{A(b, \vec{D}, \vec{F}).S}{\sigma^2} \right) - \sum_{i=1}^{N} \frac{A(b, \vec{D}, \vec{F})^2}{2\sigma^2} \tag{6.14}$$

The modified likelihood function allows estimation of $A_0$, $\vec{D}$, $\vec{F}$, and $\sigma$ in a single step. Throughout this article, we refer to this method as the joint MLE (jMLE) method.

The jMLE method assumes homoscedasticity, and any violation would result in biased estimation of the apparent diffusion coefficients. However, if homoscedasticity is fulfilled jMLE provides a better estimation of the underlying parameters. When the signal is close to noise threshold MLE estimator is also biased [226], in which case noise estimation with jMLE is better than noise estimation at each b-value. jMLE also allows less number of repeated measurements at any particular b-value and reduces the total time required for such measurements. Similar methods have been previously used in the context of DTI [235] and for diffusion kurtosis imaging (DKI) [236] but these ignored the issue of heteroscedasticity.

Recently Kristoffersen has proposed using least squares fitting to the median or mode of the PDF, while still accounting for the underlying Rician distribution [237, 238]. He found that the bi-exponential model with a rejection fraction of 2.9% fared better than all other models compared [237]. A biexponential fit is fortunately also the most commonly used method to analyze diffusion at high b-factors [119, 120, 184]. At present it seems prudent and useful to analyze the data with MLE estimate and a biexponential fit.

This study explores the possibility of performing high field (7 T), high resolution (2 mm isotropic), and high b-factor ($b_{max} = 8$ ms/$\mu$m$^2$) DWI with acceptable accuracy and precision. CSF suppression through b-filtering is compared with more common FLAIR-based CSF suppression. Two different MLE methods are compared: in the first, MLE is calculated at each b-value, and in second, jMLE, the MLE is estimated for the bi-exponential fit.

### 6.4.3 Methods

Two protocols were run - one with and another without the FLAIR-based CSF suppression [233]. The diffusion time ($\Delta$), gradient duration ($\delta$) and TE were optimized to provide maximum SNR for a maximum b-factor of 8 ms/$\mu$m$^2$. The parameters used were: $TR/TE/\Delta/\delta$ = 3000/59/121/16 ms giving 2 mm isotropic resolution. The bandwidth was 1930 Hz/pixel and the k-space was partially sampled at 5/8. Parallel imaging was not used.

The acquired region of interest extended superiorly from the dorsal surface of the corpus callosum. 16 diffusion weightings were applied in the anterior-posterior direction corresponding to y-axis. Gradients were incremented linearly in 16 steps from 6% to 95% of the maximal gradient strength such that b-values increased quadratically starting from 0.030, 0.137, 0.270, ..., 7.070, 8.0 ms/$\mu$m$^2$. Due to SAR limitations, the TR time for the FLAIR-based protocol had to be 8 s and TI was set to 1550 ms to suppress CSF. The measurement was repeated 16 times. The protocol with FLAIR took more than 34 minutes while the protocol without FLAIR took around 13 minutes.

All analysis steps were implemented in Matlab (The Mathworks Inc., Natick, MA). Diffusion decay curves were fitted to a biexponential model using the two MLE based methods described above. For the first method, data obtained from 16 repeated measurements at each b-value were used to find the MLE under the Rician distribution. The diffusional decay curves obtained after MLE were fitted to a bi-exponential decay function that minimized the weighted sum of squared residuals. The square of the standard error for the MLEs were chosen as the relative weight at each b-value. To ensure that a global minimum was found the fitting procedure was run 3 times with different starting values. For the second jMLE method, data for all b-values were used to estimate the diffusion coefficients, their volume fractions, and noise in a single pass. The data were also analyzed with b-filtering, such that b-factors less than 0.5 ms/$\mu$m$^2$ were removed from the fit.

### 6.4.4 Results

The merit of STEAM-EPI can be easily observed in the single shot DW images shown in Fig. 6.12. While diffusion weighting decreases the signal as expected, these single-shot images have a useful SNR at high b-factors. It is not uncommon for suppressed fat signal to show up at high b-factors, however no such artifacts are visible in these images. Fig. 6.13 shows MLE of the signal $S$ (top row), and the standard deviation of noise $\sigma$ (bottom row). Spatial heterogeneity is evident in the noise maps (Fig. 6.13, bottom row). The noise levels are comparable in WM and GM but elevated for CSF. For low b-value images (Fig 6.13 E,F), voxels neighboring CSF are also found to have higher noise levels. At a glance, it might seem that the higher noise level in CSF contaminated

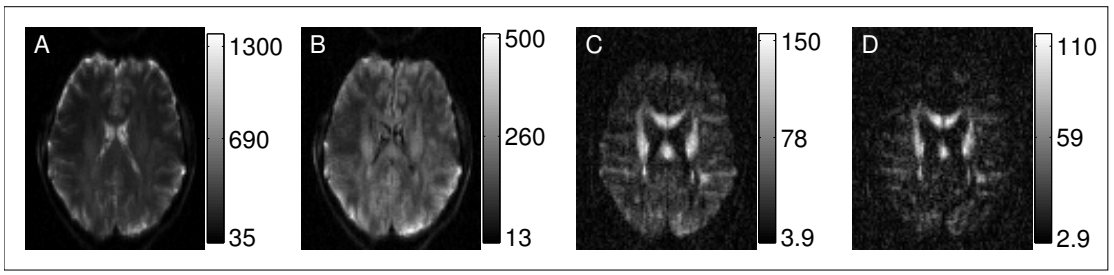



FIGURE 6.12: Single-shot images at increasing b-values (left to right) acquired with DW-STEAM-EPI. The b-values are 0.03, 0.5, 3.8 and 8 ms/$\mu$m$^2$. Diffusion weighting was applied in the anterior-posterior direction for all acquisitions.

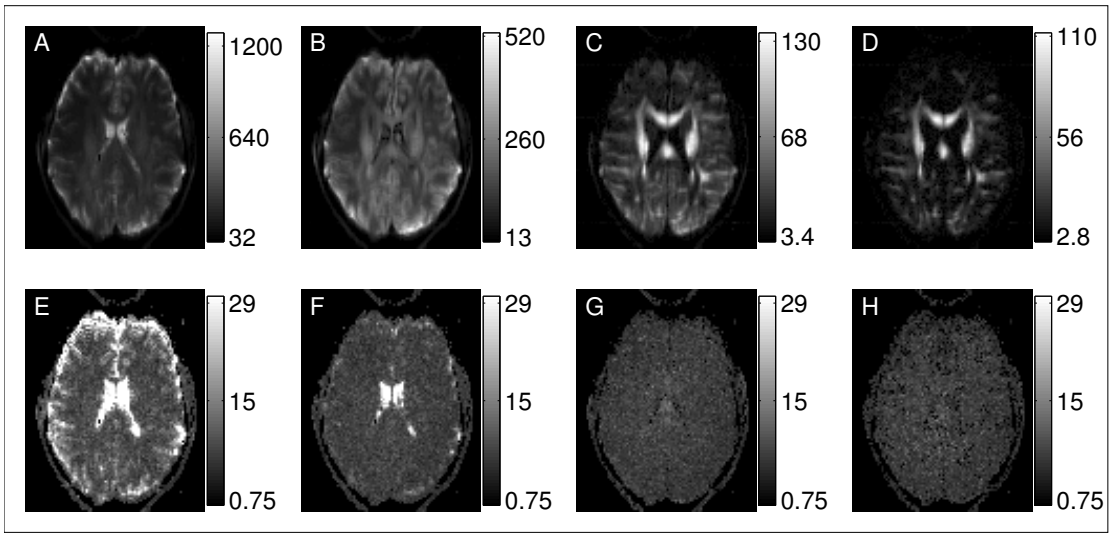



FIGURE 6.13: Maximum likelihood estimate of the signal (top row) and standard deviation of noise (bottom row) obtained with 16 repetitions. The b-values are in increasing order from left to right (0.03, 0.5, 3.8 and 8 ms/$\mu$m$^2$).

voxels is related to higher signal. However, we find CSF contaminated voxels to have higher noise level even for the FLAIR-based protocol in which these voxels have lower signal than surrounding tissue.

Fig. 6.14 shows MLE of the signal (top row), and standard deviation of noise (bottom row) obtained for the same slice but with FLAIR based CSF suppression protocol. In these images, notwithstanding a drastic reduction in CSF contribution to the signal, CSF voxels still show elevated noise levels. As the b-value is increased, the CSF contribution to signal is rapidly diminished and the noise maps look homogeneous.

Fig. 6.15 summarizes this information in 2D intensity maps of mean estimated Rician noise at different b-values and signal intensities. For the data obtained with FLAIR, there is very little correlation: estimated Rician noise depends neither on the signal intensity or nor on the b-value (Fig. 6.15A). For the data obtained without FLAIR (Fig. 6.15B), at low b-values the noise decreases with increasing b-factors. However,

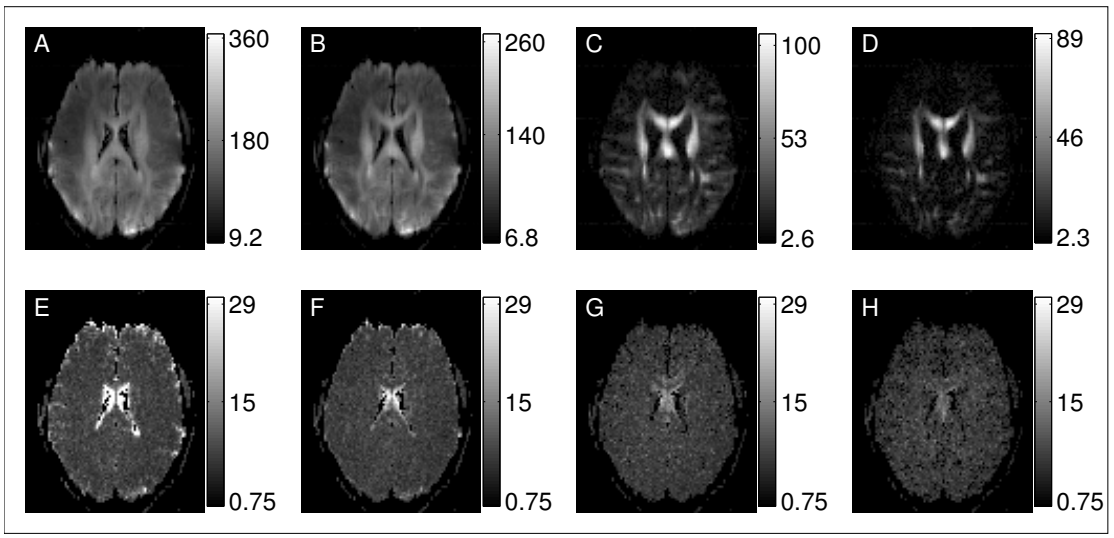



FIGURE 6.14: Maximum likelihood estimate of the signal (top row) and standard deviation of noise (bottom row) for images acquired with FLAIR protocol. Voxels containing CSF contained very little signal but elevated noise levels compared to surrounding tissue.The b-values are in increasing order from left to right (0.03, 0.5, 3.8 and 8 ms/$\mu$m$^2$)

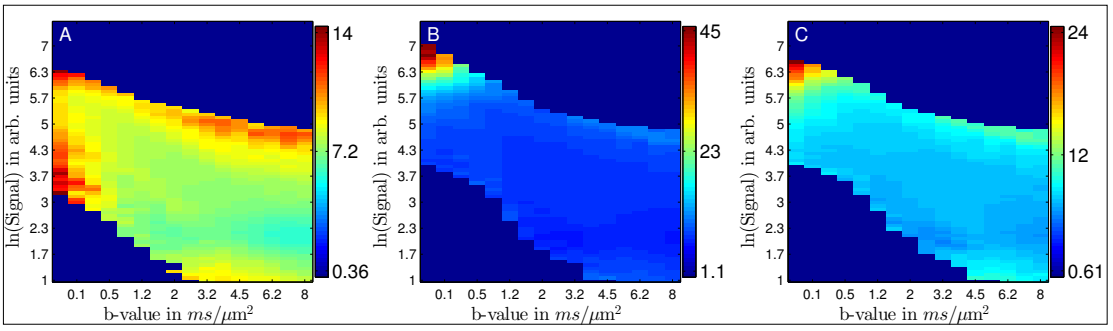



FIGURE 6.15: Intensity maps of the mean Rician noise estimated at different b-values and signal intensities. On the left (A) is obtained from data acquired with FLAIR based CSF suppression and middle (B) without FLAIR. Image C on the right is obtained with same data as B, but only for voxels with CSF-ratio $< 3$. Note that the x-axis the scale is not linear but monotonically increasing for 16 b-values measured.

this dependence mostly comes from voxels with CSF partial voluming. To assess the influence of CSF in physiological noise, we introduce a term 'CSF-ratio' defined as the inverse of the ratio between the signals with and without FLAIR protocol, both acquired at the lowest b-value (0.03 ms/$\mu$m$^2$). Since FLAIR suppresses around 90% of the CSF-signal and around 30% of the signal from WM and GM, a voxel with a CSF-ratio of 3 contains around 60% CSF in it. Fig. 6.15C shows expectation maps from the same data set as Fig. 6.15B but only for voxels with CSF-ratio less than 3. At low b-values there exists some noise elevation but this rapidly diminishes with increasing b-factor. For all b-values greater than 0.5 ms/$\mu$m$^2$, the mean signal varied from 250 to 30 whereas

the noise had a small range between 7.5 and 9. The maps clearly reveal that variation in noise is only an issue at low b-values where physiological noise manifests itself in voxels containing CSF or CSF partial volumes. Removal of variations in noise with increasing b-values is encouraging: it suggests that CSF suppression with b-filtering could help mitigate CSF partial voluming and also remove heteroscedasticity across b-values. Demonstration of homoscedasticity along the diffusional decay curve also paves the way to pursue our predilection in using jMLE with b-filtering.

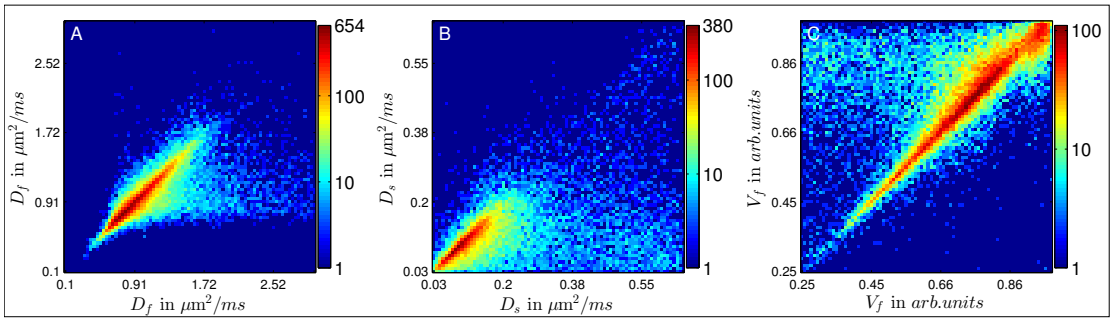



FIGURE 6.16: Correlation maps of $D_f$ (left, A), $D_s$, (middle, B), and $V_f$ (right, C) obtained using two different biexponential fitting methods. On x-axis is the jMLE fitting where Rician estimate as well as biexponential fit is done in a single run, on y is the fit done separately after maximum likelihood estimate of the signal at each b-value. The intensities are shown in logarithmic scale and represent number of voxels. The correlation is very high, only very few voxels show deviation in the values between the two methods.

## Fitting

The data obtain using FLAIR-based CSF-suppression was used to compare the two fitting procedures. Since heteroscedasticity is not severe and there is little residual CSF partial voluming in this data, there should be little difference between the estimates obtained from two biexponential fitting procedures. Fig. 6.16 shows correlation between the parameters estimated from these two methods. As expected, for most voxels, the two methods provide almost identical results. This confirming that heteroscedasticity is not a problem and jMLE can be used as a reliable method.

Although few, there are some voxels where differences exist. Further analysis showed that these are voxels where the 'true' signal is very low at high b-values. As mentioned before, when the true signal $A$ is close to the Rician noise $\sigma$, then ML estimator is biased [226]. This bias further propagates during biexponential fitting, providing erroneous results. This can be seen in Fig. 6.17 which shows the mean difference between the parameters obtained from the two methods as a function of the SNR at the highest b-value. Deviation between the two methods is only apparent at low SNR, and increases with lower SNR.

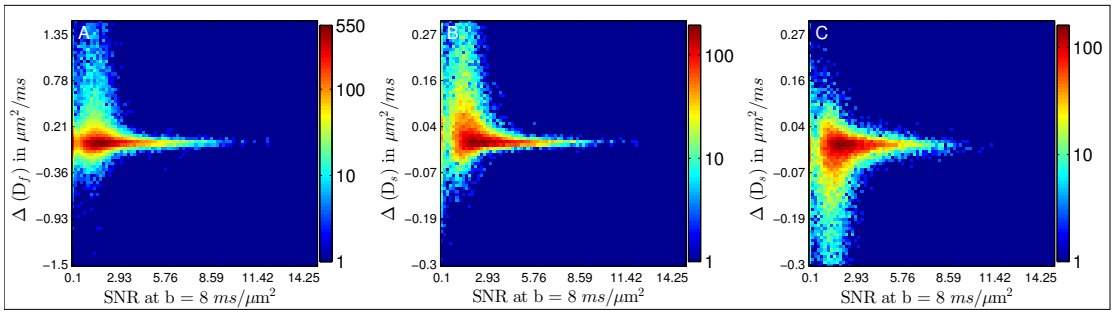



FIGURE 6.17: Difference in $D_f$ (left, A), $D_s$, (middle, B), and $V_f$ (right, C) obtained using two different biexponential fitting methods. The intensities are shown in logarithmic scale and represent number of voxels. For most part the deviation is very small but gradually increases with decreasing SNR. This deviation can be attributed to the bias in ML estimator at low SNR.

#### 6.4.4.1 b-filtering

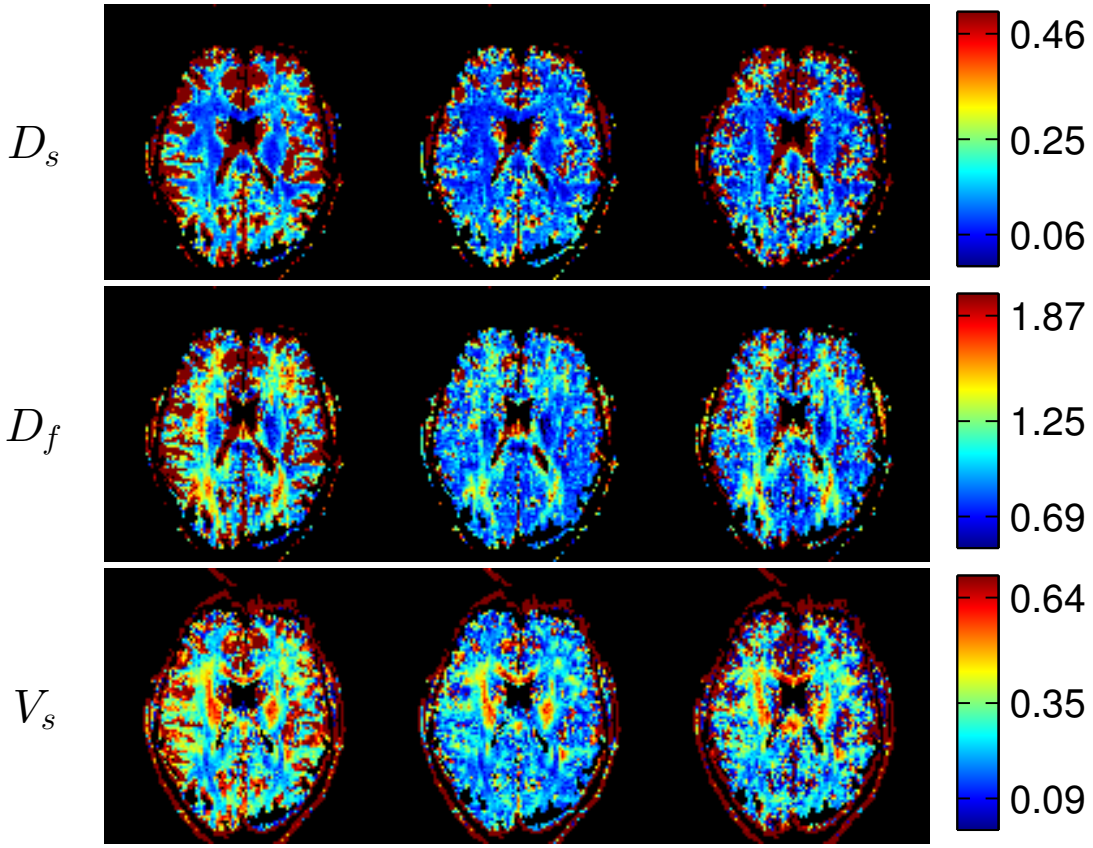



FIGURE 6.18: Results for the biexponential fit parameters, before b-filter (left), after b-filter (middle) and with FLAIR (right). The top row is the slow diffusion component ($D_s$), the fast component ($D_f$) is in the second row and the third row shows volume fraction of the slow component ($V_s$). Without any filtering (first column), CSF partial voluming has affected many voxels. B-filtering (second column) largely mitigates this issue and the maps are comparable to those produced with FLAIR based filter (third column).

Figs. 6.18 shows maps of ADCs for the slow component $D_s$, (top row), the fast component, $D_f$ (middle row) and volume fraction of the slow component $V_s$ (bottom row) obtained without any filtering (first column), with b-filtering (second column) and with FLAIR based CSF suppression (third column). Fig. 6.19 shows CSF-ratio map for a representative slice. It was found that for voxels with very low CSF contribution (CSF-ratio $< 2$), the biexponential fits were similar (as expected) with or without FLAIR. For these voxels, b-filtering slightly underestimates all three: the $D_f$ by almost 5%, $D_s$ by almost 10% and the $V_s$ by 3%. However, as CSF-ratio increases, the estimated values for $D_f$ and $D_s$ obtained without CSF suppression are much larger. For these voxels b-filtering produces similar results to those obtained with FLAIR-based CSF-suppression. In the absence of b-filtering, overestimation due to CSF partial voluming were found to be the most severe for $D_f$.

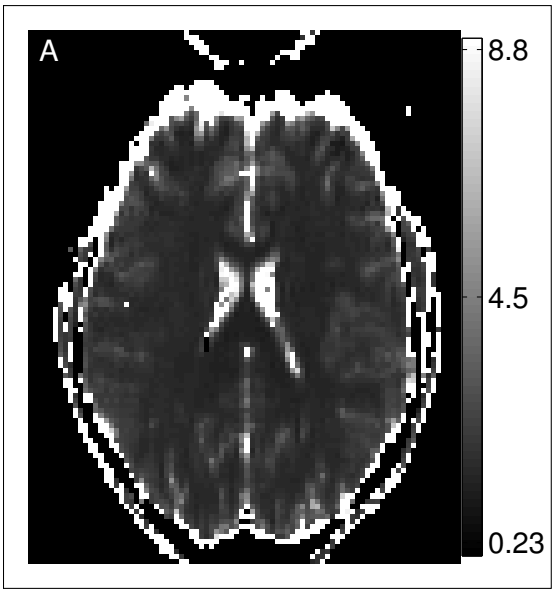



FIGURE 6.19: CSF-ratio map obtained by dividing signal intensities obtained at the b-value of 0.03 ms/$\mu$m$^2$ with another CSF-suppressed image at the same b-value. To show good contrast for gray and white matter regions the intensity threshold is limited to 10. All voxels with the ratio greater than 10 contain all their signal from CSF. As expected WM shows very little CSF contamination.

### Repeated Measurements

To obtain confidence intervals of our estimates, we computed the Hessian of the log-likelihood for the MLEs of each parameter. The negative inverse of this Hessian gave cumulative distribution function (CDF) of the estimated parameters. Maximum likelihood estimates are asymptotically normal - the confidence intervals are estimated by computing the inverse of the CDF under a normal distribution. Fig 6.20 shows the 95% confidence intervals for $D_f$, $D_s$ and $V_f$ respectively. As revealed in these images

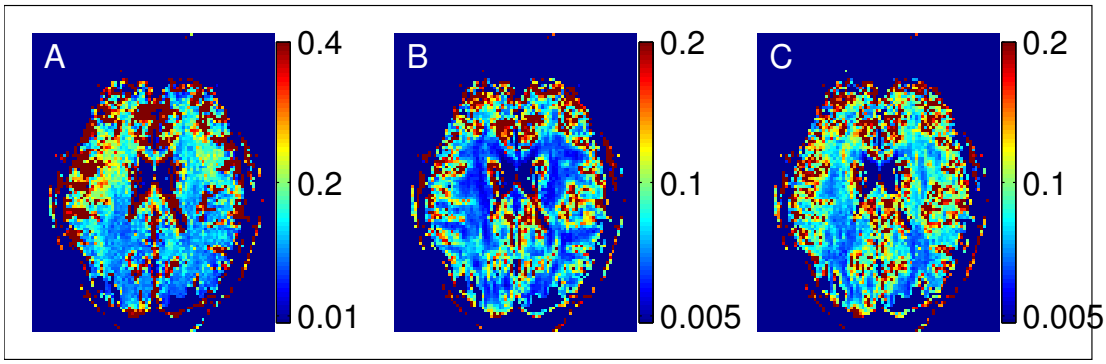



FIGURE 6.20: Maps of 95% confidence interval for $D_f$ (left, A) , $D_s$ (mid,B) and $V_s$ (right, C) respectively. Smaller confidence interval mean higher confidence in the parameter estimates. Regions in the white matter show higher confidence compared to regions in gray matter, particularly close to the CSF. This could either be due to partial voluming or the insufficiency of the model.

the confidence intervals are rather narrow, indicating robust repeatability of these measurements. The median of the 95% confidence for $D_f$ was 0.19 for all voxels, but for most WM voxels with CSF-ratio $< 2$, the mean 95% confidence is only 0.14. It is clear that the variability is higher in gray matter, particularly in regions neighboring CSF. This increase in uncertainty could be due to both higher noise in those regions or the insufficiency of the model itself. Nevertheless, the model holds well for voxels in white matter.

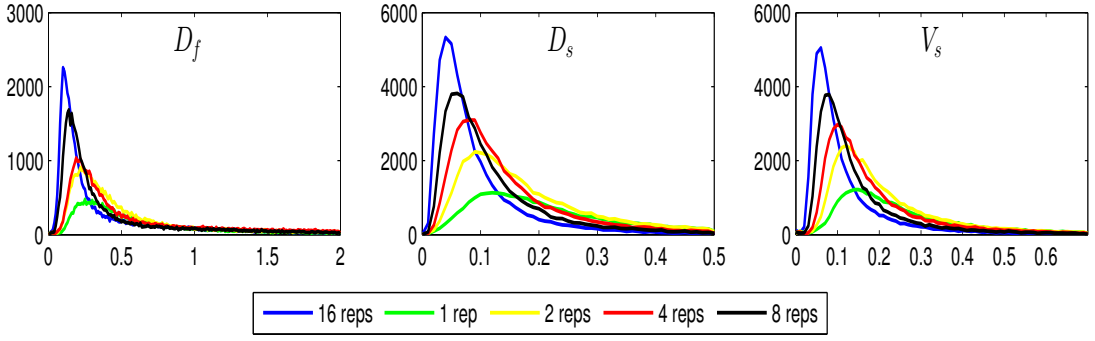



FIGURE 6.21: Histograms of 95% confidence interval for $D_f$ (left, A), $D_s$ (mid,B) and $V_f$ (right, C) at different number of repeated measurements. Smaller confidence interval mean higher confidence in the parameter estimates.

While 256 measurements for each diffusion decay curve allow the parameters to be estimated with high confidence, it might be expensive, time consuming or even superfluous to acquire so much data. To assess the possibility of using fewer data points, as a first approach, we selected a random subset of data for each b-value and computed the 95% confidence intervals for each case. Fig 6.21 shows the histogram of these confidence intervals for different number of repeated measurements ( 1, 2, 4, 8, 12). Increasing

number of repeated measurements indeed increases confidence in the estimates, making the histograms sharper. Even with 4 averages the median 95% confidence interval for $D_f$ and $D_s$ is around 0.3 and 0.1. This is acceptable and comparable with confidence intervals obtained for DTI parameters acquired for DTI parameters acquired more conventionally [239].

### 6.4.5 Discussion

This study explored an acquisition and analysis strategy for time-efficient high b-factor DWI. While most of the methods described here have been previously reported, several of them have not been widely used, and others have been overlooked altogether. STEAM-EPI, for instance, has been very rarely used for DWI in the brain. However, it is clear that it gains relevance with increasing field strength. The choice between spin-echo or stimulated-echo depends on field strength, desired b-value and the gradient strength. Data with high b-values and reasonable SNR are much easier to achieve with STEAM-EPI. Furthermore, compared with spin-echo methods, this sequence has a reduced SAR giving it further usefulness at high fields. If similar RF pulses are used, the SAR in stimulated echo is 40% less than in spin echo and 66% less than in twice-refocused spin echo. With the available gradients, stimulated echo is evidently more suitable for high b-factor DWI. The clarity of single shot images are proof of its quality. To my knowledge the images obtained with this method are so far the best in-vivo human brain images yet obtained at high b-factors and 2 mm isotropic resolution. It is worth mentioning that an isotropic voxel size of 2 mm contains almost half the volume of a 2.5 mm isotropic voxel (for comparison please refer to [119, 137]).

Another noteworthy observation from this study is in relation to physiological noise. It has been widely reported that physiological noise in DWI depends on signal magnitude. This study shows that the increase in noise baseline ceases at a b-value of around 0.5 ms/$\mu$m$^2$, suggesting that the origin of this elevation is mainly CSF. Apart from DWI, physiological noise is a nuisance in other MR modalities and correction schemes have been suggested [237, 240–243] Since an appropriate level of diffusion weighting successfully eliminates the noise due to CSF partial voluming, it can also be added to other imaging modalities such as fMRI. Although b-filtering reduces signal, the reduced CSF contribution and thus less physiological noise may still result in enhanced sensitivity in BOLD imaging. Previous studies using lower b-values ( less than 0.1 ms/$\mu$m$^2$) to remove vessel contribution have also reported on increased sensitivity [244]. We also find that while FLAIR based suppression still retains the elevated noise baseline, b-filtering gets rid of it completely. This noise is caused by flow effect during the inversion time used in FLAIR, which is not a problem for b-filtering. B-filtering also has the added benefit of adding no SAR contribution. For high field MR, b-filtering can replace FLAIR based

suppression for all MR contrasts. With the increased sensitivity it can provide it could also find use in regular BOLD imaging.

While b-filtering is easy to include in other MR modalities, its use in DWI itself is not intuitive. When CSF partial voluming is not an issue, it is best to use all of the diffusion decay curves to estimate the diffusion parameters. A better approach is to use b-filtering where necessary. It is easy to obtain one image at low b-value using FLAIR based suppression, and to use this image to calculate the CSF-ratio. B-filtering can then be applied in voxels where the ratio is high, by excluding the low b-value data from the calculation of the diffusion parameters. Prudent use of b-filtering method can help mitigate problems due to CSF contamination without time and SNR penalty.

Another important issue addressed here is combining Rician distribution in the signal amplitude with the model for diffusion decay. This study used a biexponential fit but such the jMLE method can be performed with any other model. When heteroscedasticity is not an issue, a direct approach is better in estimating parameters than the two-step process in which MLE is obtained in the first step and the model is assessed in the second. If only four repetitions are needed to provide reliable estimates, then scanning in each diffusion direction could be done in around four minutes, allowing scanning with more directions in a reasonable time. Furthermore, since repeated measurements are not necessary for a single b-value, a model-dependent better sampling along the decay curve can be used. The only drawback of jMLE technique is its inability to compare between models. Parameters estimated from the jMLE fit depend on the model and using them to compare between models is not appropriate. In this case a better solution is to obtain noise threshold from low b-value data and use it as a given quantity while estimating MLE at low b-values. This method has been applied in the next measurement.

## Conclusion

This study explored different ways to efficiently obtain and analyze high b-value data. We have also shown that physiological noise mostly arises from the CSF contribution to the signal, and can be effectively mitigated with sufficient diffusion-weighting.

For an efficient data collection, it is better to obtain more repetitions for data at high b-values that have a low SNR. However, due to physiological noise, the low b-value data had much more variance than a slightly higher b-value data. Fortunately, since diffusion decay curves are exponential, it is possible to remove such low b-value data from diffusion measurements without losing information. It was also found that such process actually helps to remove partial voluming from CSF.

The goal in the next experiment is to measure diffusion decay curves at long diffusion times and higher b-values.

## 6.5 DWI at high b-factors and long diffusion times

The non-monoexponential decay of diffusion signal at high b-factor is an unequivocal reflection of heterogeneity in the tissue structure. Nonetheless, the diffusion process in itself is a homogenizing phenomenon: if the transit time is long enough, the diffusing particles that start from a given position can be expected to be found in all clefts and corners of the tissue. At such extremely long times, the probability of finding a particle in any region is proportional to the fraction of total volume that the region occupies. In this limit, the central limit theorem applies and the resulting diffusion propagator is effectively be a Gaussian with diffusion coefficient $D_{t\to\infty}$ that is equal to the weighted mean of the diffusion coefficients in each region (Sec. 2.6.2).

It is axiomatically true that the diffusion coefficient is time independent once it reaches the long time limit, however, it is not trivial to estimate when this long time limit is attained. On one hand, the long time limit is extended due to structural complexity and on the other, it is shortened by increased diffusivity of the probing particle. In bulk liquid, the ballistic trajectory of particles, that occur in a time scale of nanoseconds quickly become Gaussian [245]. But if the medium is composed of several heterogeneous structure with low permeability the long time limit could be well beyond experimentally observable times. If the probing particle is a highly diffusive gas, then the heterogeneous compartments are quickly probed and the long time limit is attained faster. This interplay between diffusivity and structural complexity makes time dependent diffusion studies interesting and challenging. In the DTI measurement (Section. 6.2), it was found that the long time limit is not quite reached even for the diffusion time of 2 seconds, but some indication of approach towards Gaussianity was seen.

Using diffusion MRI in WM, very few studies have dealt with time dependence of the diffusion decay curves. In the simulations using anatomically correct boundaries, Chin et al. [246] concluded that exchange between intra and extra axonal water should result in a monoexponential diffusion decay curves for all Δ greater than 300 ms. Similarly, another Monte Carlo simulation by Fieremans et al. [247] performed for two-compartment exchange model in WM model showed that kurtosis, the deviation from Gaussian diffusion, attained a value close to zero when the simulated diffusion time was around 15 times the exchange time. The maximum reported exchange time in that study was 20 ms which again corresponds to the supposed long time limit of 300 ms.

Experimental findings however suggest otherwise. Measurements performed on excised spinal cord for diffusion times between 50 and 200 ms did not show any sign of the decay curves moving towards mono-exponential decay [138]. On the contrary, it suggested some restriction on the order of 8 to 10 $\mu$m. Similarly, for brain WM *in vivo* Nilsson et al. measured diffusion perpendicular to the fiber direction in corticospinal tract and reported clearly non-monoexponential decay for diffusion times between 64 and

256 ms [248]. They reported that the data could not be sufficiently explained without considering exchange between the two components but it was clear that long time limit had not been reached at 256 ms.

To our knowledge Nilsson et al.'s is the only study reported on human WM *in vivo* at such long times [248]. Their study reveals interesting facts about diffusion process and pose questions about the exchange mechanism between heterogeneous regions in WM. Including exchange to the two compartment model they obtained mean exchange time of about 300 ms. They also reported that ignoring exchange the estimated ADC of both the fast and the slow pool did not change with $\Delta$. However, if the exchange time is around 300 ms, then how could diffusion decay curves obtained at $\Delta_{min} = 64$ ms, and $\Delta_{max} = 256$ ms be very similar to each other? Additionally, the relative volume fraction was also found to be independent of $\Delta$ which suggests that there ought to be very little exchange at those times.

The availability of 7 T system and strong gradient coils allow us to perform high b-factor measurements with long diffusion times. The primary goal of this work is to obtain information about the behavior of water diffusion at large range of diffusion times and follow the evolution of estimated parameters. Since diffusion decay is less steep when measured radially to the white matter fibers, it is easier to obtain respectable SNR even for high b-factors. In this study the measurements performed in y-direction which is perpendicular to both the corpus callosum and the corticospinal tract that are the two major coherent fiber bundles in brain. We have explored maximum diffusion time of 1500 ms which is 5 times the exchange time reported by Nilsson et al. [248]. The features of the decay curves were first separately examined for each diffusion time, then Kärger equations for exchange was applied and then a linear fit between signal and $1/\sqrt{b}$ was performed. The last fit was used to distinguish if the major contributor to the observed non-monoexponential behavior was due to glial processes or tubular water.

### 6.5.1 Methods

Three volunteers were measured. Images were acquired with 2 mm isotropic resolution using 2D single shot DW-STEAM-EPI. The acquisition bandwidth was 1930 Hz/pixel and the k-space was partially sampled at 5/8. The diffusion times ($\Delta$) used were 75, 150, 300, 675, 1500 ms. For all the measurement times the echo time varied between 33 and 48 ms. To accommodate many slices for long $\Delta$ measurements, the repetition time (TR) had to be varied between 4.5 and 6.5 seconds. However, $T_1$ of the tissue is around 1500 ms and difference in TR times does not affect the results. Even for these relatively long TR we could only acquire 4 slice for the longest diffusion time of 1500 ms.

14 diffusion weightings were applied in anterior-posterior direction corresponding to y-axis. Gradients were incremented linearly in such that b-values increased quadratically.

To improve SNR, images with high b-factors were repeated more often than images with low b-factors. For instance, b-value of 0.7 ms/$\mu$m$^2$ was repeated 6 times whereas 15 ms/$\mu$m$^2$ was repeated 15 times. A total of at least 126 scans were performed for each Δ except for 1500 ms where 168 scans were performed. To select coherent single fiber population a DTI measurement was additionally performed in the same session. Only the single fiber WM voxels selected by the DTI measurement were investigated.

The MLE was performed independently at low b-value data and for the high b-value data, the noise from the low b-value data was taken as a known value. Initially bi-exponential fits were performed for all curves obtained at different Δs. These results obtained from these individual fits are reported and role of exchange is discussed. The decay curves from all Δs were also combined and fitted with a two-component fit including exchange using Kärger's exchange formulation.

To discriminate between possible sources of the non-monoexponential decay, the signal at high b-factors was also fitted to a $1/\sqrt{b}$ dependence for each Δ.

### 6.5.2 Results

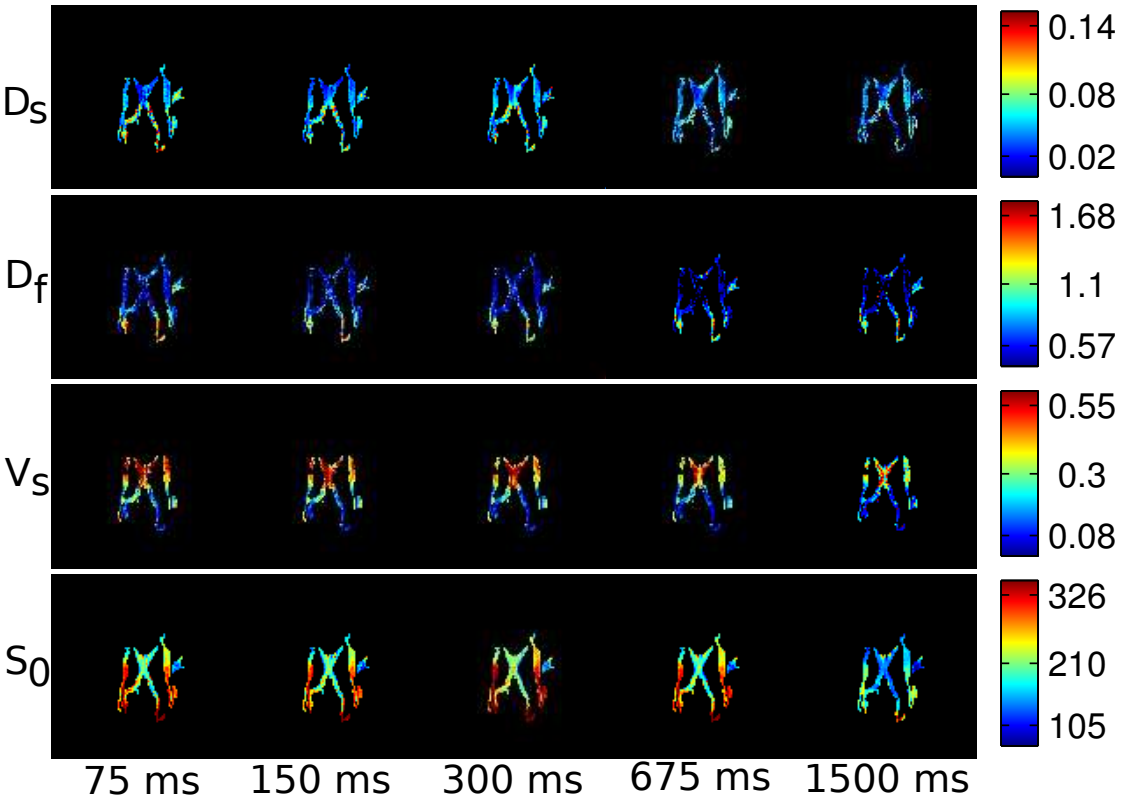



FIGURE 6.22: Maps of $D_s$ (top row) , $D_f$ (second row) , $V_s$ (third row) and $S_0$ (bottom row) obtained with biexponential fits for at different diffusion times. The diffusion times are in ascending order from left to right 75 ms, 150 ms, 300 ms, 675 ms, and 1500 ms.

For each Δ, the curves could be described well with biexponential fits. Fig. 6.22 shows typical maps of fast ADC ($D_f$), slow ADC ($D_s$) and volume fraction of $D_f$ ($V_s$) obtained for different Δs. For the fast diffusion coefficient $D_f$ and the slow diffusion coefficient $D_s$, the maps are very similar for different Δs although the longest time of 1500 ms was 20 times longer than the shortest time of 75 ms.

The results were quite close to expectations that could be drawn from the DTI measurements reported in section 6.2 and the high b-factor measurements in section 6.3. The biexponential fit from the previous measurement had shown that the $D_f$ and $D_s$ differed by an order of magnitude. The DTI data had shown very small decrease in ADC with increasing diffusion time: to get the same ADC, a huge drop in $D_f$ could not be expected. The $D_s$ was already close to zero, any small change would not be significant to affect ADC.

As expected, the fast diffusion coefficient $D_f$ decreased slightly with increasing Δ and the slow diffusion coefficient $D_s$ showed very small time dependence. But, the relative volume fractions showed a clear dependence with diffusion time. Fig. 6.23 shows $V_s$ for all fiber angles and different Δs measured. In the parallel direction we see that there is a small drop from 0.15 at 75 ms to around 0.1 at 1500 ms. Although $V_s$ is fairly small, still even at 1500 ms MwB fit was not sufficient to describe the diffusion decay curves and biexponential fit had to be performed. The change in $V_s$ is more prominent with increasing angle between the gradient direction and the fiber orientation. For fibers orientated close to 90° to the y-axis, $V_s$ is about 0.5 at 75 ms and drops to around 0.3 at 1500 ms. A decreasing $V_s$ with increasing Δ is a strong indication that exchange has played a role in determining the diffusion decay curves.

#### 6.5.2.1 Exchange

Encouraged by a clear indication of exchange, the data from all Δs were combined and fitted with a two-component fit including exchange. For comparison, a simple biexponential fit was also performed for this combined data. Exchange equations do not account for decrease in the ADC due to restrictions. Therefore, theoretically applying Kärger model to our data might require additional parameters to account for decreasing ADC with increasing Δ. However, given the decrease in ADC are very small, adding such parameter would unnecessarily complicate the model without adding much value in understanding of the data.

As expected, the biexponential fit could not fit the combined data. However adding the exchange parameter considerably improved the fitting procedure. Fig. 6.24 shows the reduced $\chi^2$ values obtained from the fits with and without including exchange. It is observed that including exchange is only relevant when the measured direction is perpendicular to the fiber orientation. When the measured direction is perpendicular

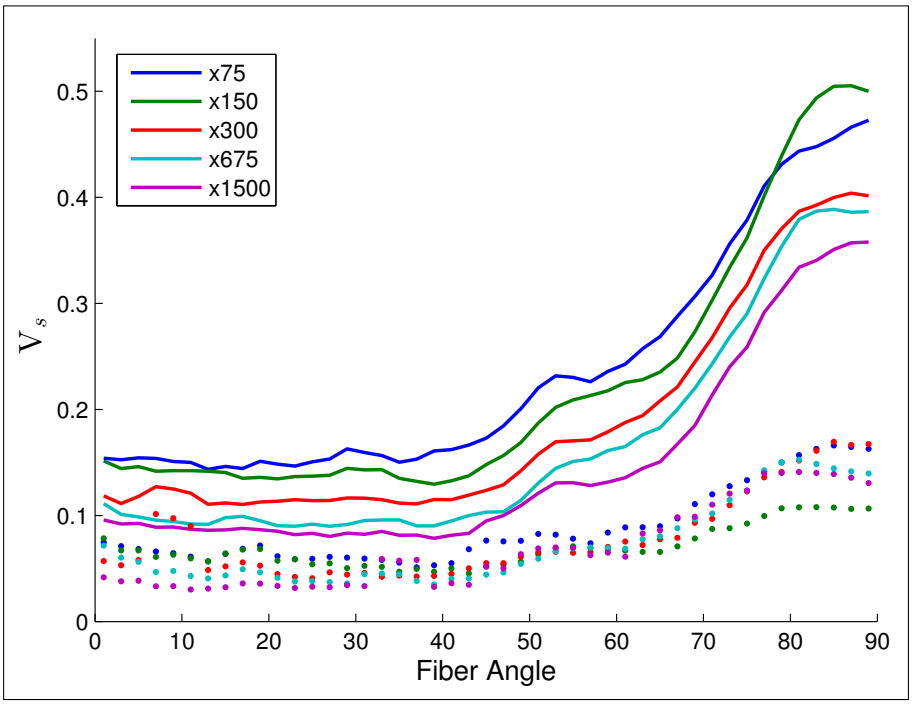



FIGURE 6.23: Volume fraction of the slow component (solid lines) against diffusion time for different fiber orientations. As the diffusion time is increased the volume fraction decreases steadily. However, a significant fraction still remains even at the longest diffusion times. The dotted lines are standard deviation of the data.

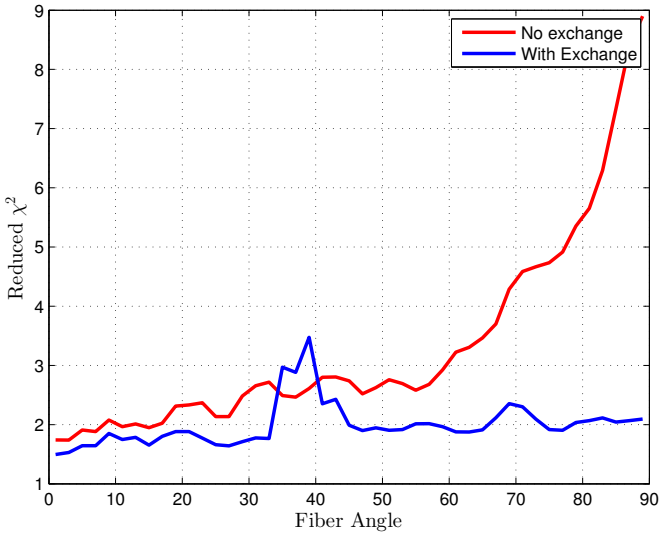



FIGURE 6.24: Reduced $\chi^2$ values with and without exchange.

to the fiber orientation, inclusion of an extra exchange parameter results in a four-fold improvement in the fit. But in the parallel direction, there is almost no improvement and the biexponential fit seems fairly sufficient to describe the decay curves for all directions. This is also reflected in the exchange value obtained from the fit. Fig. 6.25a is the scatter plot of the exchange parameter against the angle made by the WM fiber

with the measured y-axis. For small angles the estimated value almost always hit the upper limit of 15000 ms. In parallel direction the biexponential fit is already good and adding exchanging will not contribute much to the goodness of the fit(see Fig.6.24). Even when the upper limit was raised to $10^6$ ms, many voxels in parallel direction still hit the upper limit. Since the maximum diffusion time in this study was 1500 ms, higher values are beyond meaningful estimation. When the angle is more than 50° the exchange times get smaller and within 90° it reach a mean value of around 1500 ms and a median of about 1250 ms. Fig. 6.25b shows the histogram of exchange times when the diffusion direction is greater than 80° further from the fiber direction. Since there is a long tail in the exchange time distribution, the median of 1250 ms is probably a more meaningful number than the mean.

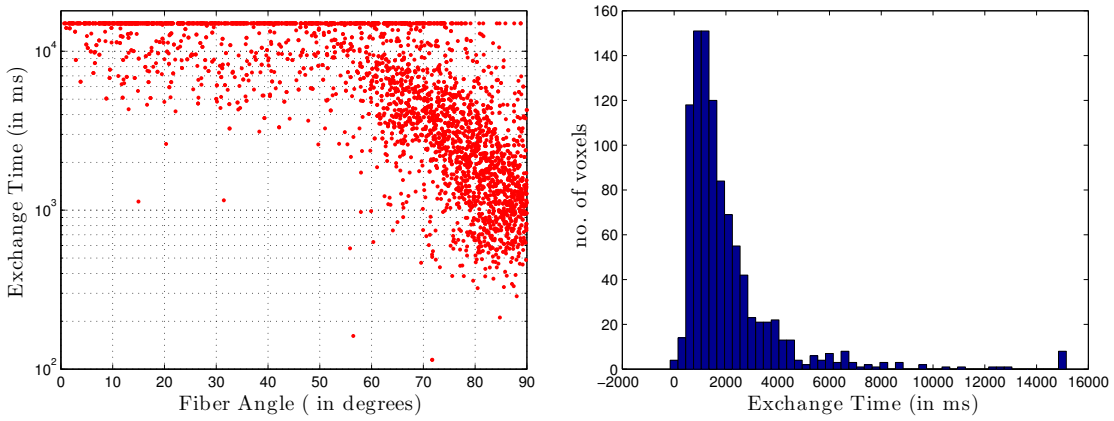



FIGURE 6.25: Left: Scatter plot of exchange time of the fast component against angle made by the measurement direction with the fiber direction. The exchange parameter is only relevant when the measured direction is close to being perpendicular to the fiber direction.Right: histograms of the exchange time for fibers lying almost perpendicular to the measured direction.

An interesting feature in Figs. 6.23, 6.24,and 6.25 is the transition around 50° or perhaps 45°. It seems that at this point the biexponential fitting switches from one kind of two compartment system to another. Until this point $V_s$ is relatively flat but then it starts accounting for the increased volume fraction that is restricted in the perpendicular direction. Similarly, when all the data were combined (Fig.6.24), the biexponential fit still performs fairly well until around the transition point and then it fails to fit the curves and exchange has to be included. The estimated exchange values are then relevant only after 50° angle.

This transition around 45° is when the axonal contribution to the slow component has to be accounted for. Taking the intra-axonal ADC in the parallel direction to be around 1 $\mu m^2$/ms, its ADC at 45° would be 0.5. Therefore, at a b-value of of 5 ms/$\mu m^2$, the signal contribution from axonal water would already be very small. These high b-value

data can be used to differentiate between the two possible compartments that could have contributed to the slow component (i) tubular water and (ii) water in glial processes.

### 6.5.3 High b-factor signal as function of $1/\sqrt{b}$

The major feature of glial water is its isotropy. If the slow water arising in parallel direction is mostly due to glial processes, then the decay curve at high b-values should show a $1/\sqrt{b}$ dependence independent of the small variation in angle. In this case, for high b-factor (b $>$ 5 ms/$\mu$m$^2$) measurements parallel to the fibers (up to around 45°), a linear fit between $1/\sqrt{b}$ and the signal should produce a slope that is relatively independent of the measurement direction. From Eq. 4.2, it is clear that when diffusion perpendicular to the processes is close to zero, the slope of this fit would be equal to $V_{gp}\sqrt{\pi}/2\sqrt{D_gp}$, where $V_{gp}$ is the volume fraction of the glial processes and $D_{gp}$ is the diffusion coefficient of these processes in the direction along their main axis. If the contribution is mostly from tubular water, then such a fitting procedure is not meaningful and should show a clear direction dependence.

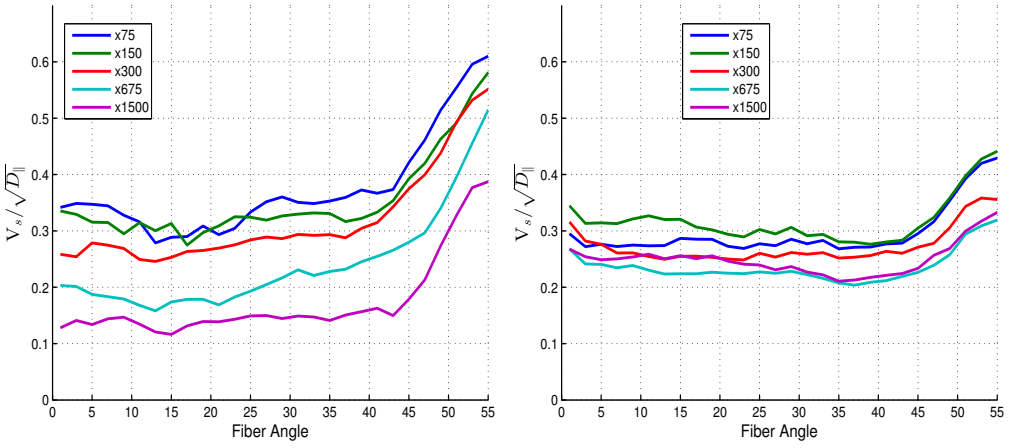



FIGURE 6.26: Direction dependence of the slope of high b-factor signal against inverse square root of b-values. The fit is only meaningful for measurements near parallel direction where axonal water does not contribute to the signal. There seems to be no direction dependence between 0 and 50°. This suggests that the signal at high b-factors mostly arises from a isotropic distribution of glial processes. The y-axis in this figure corresponds to $V_s/\sqrt{D_\parallel}$, where $V_s$ is the volume fraction of the glial processes and $D_\parallel$ is the ADC in the direction parallel to each process. On the right is the result from a linear fit. On the left is the result when the fit was forced to pass through the origin.

Fig. 6.26 shows the slope of the curve obtained by fitting the signal against $1/\sqrt{b}$. On the left side is the shows the slope obtained by fitting the signal for b-factors greater that 6 ms/$\mu$m$^2$ to $\sqrt{\pi}/2\sqrt{b}$ (Eq. 4.3). Since the processes only occupy some volume fraction of the total signal the value of the y-axis is the ratio between their volume fraction ($V_s = V_{gp}$) and $\sqrt{D_\parallel}$. Since an increase in $D_\parallel$ is not possible with increasing $\Delta$, $V_{gp}$

must be then assumed to be decreasing with increasing diffusion times. This is plausible due to exchange at such long times.

On the right side of Fig. 6.26 the same fit is shown, but this time the fit is strictly forced to pass through the origin so that Eq. 4.3 is forced. In this case the values obtained between different Δs are closer together.

Nevertheless, it is clear from these curves that indeed the signal can be expressed as a function of inverse root of b-values. For a long range of angles, the values are independent of the direction measured. This strongly suggests that the signal is indeed mostly contributed by water molecules residing in the processes of the glial cells.

## Discussion

This study shows that a monoexponential fit is not sufficient even at very long diffusion times, neither in parallel nor in perpendicular direction. This finding is different to that mentioned by Nilsson et al. where they reported a monoexponential fit for diffusion in parallel direction. They mentioned that this could mainly be due to SNR limitations which is probably the case.

Unfortunately, the necessity of the non-monoexponential diffusion decay curves in the parallel direction has been grossly overlooked by many present day modeling schemes. In light of a present result that shows a persistent non-monoexponential behavior across all diffusion times, the information obtained from complex models that omit an essential feature need to carefully assessed.

The most interesting trend is reflected in $V_s$ which continuously decreases with increasing Δ. When $V_s$ is close to unity or zero then the data is basically monoexponential. Thus, a move towards unity showed that the diffusion pools are indeed mixing at longer Δ but the mixing is not as fast as those used in simulation studies[246, 247].

After applying the Kärger model for exchange we find that the exchange is only meaningful when diffusion direction was perpendicular to the fiber direction. In the perpendicular direction there is significant exchange between the 'slow' and the 'fast' water pools and the median exchange time was around 1250 ms. However, the 'slow' pool found in the parallel direction has a rather long residence time. This pool does not exchange much and showed an exchange rate less than that of the axonal water.

Water trapped in some long microtubules might have long residence times. But, these are very rare and microtubules with 13 nm inner diameter could not possibly hold enough water to account for such a significant slow water contribution at these extremely long times. Therefore, the water in the tubules are probably not the primary contributor for the slow diffusion component observed in the parallel direction.

The only option available was the water residing in the glial processes. Glial processes are long and narrower than the axons. If the glial membranes are not so permeable to

water then the it is possible for the water there to reside within the processes - even longer than axonal water. After all, axonal water has exchange sites at nodes of Ranvier that appear every 50-100$\mu m$ whereas the glial processes, both for oligodendrocytes and astrocytes, have exchange site only at their end feet or the cell soma [249, 250][1]

Given the more of less isotropy of the glial processes, it follows that the 'slow' pool in the perpendicular direction is composed of two different water pools: axonal water, and glial processes. Most of the slow water pool in the perpendicular direction should be the intra-axonal water.

The signal dependence at high b-factors also provide additional support to the conjecture that the glial processes are indeed the primary contributor to the slow diffusion component in the parallel direction. For measurements significantly away from the parallel direction, as long as the axonal water contributed very little to the signal, the slope of the signal again $1/\sqrt{b}$ was found to be independent of the angle. The relative independence of the slope for different measurement angles is characteristic of a isotropic system.

Given that water in the glial processes is difficult to separate from the axonal water in the perpendicular direction, the exchange time between intra and extra-axonal water is perhaps slightly overestimated. If 10% volume fraction of the glial processes is accounted for, the exchange time between the intra and extra-axonal water could be down to around a 900 ms.

### 6.5.4 Conclusion

To my knowledge, this is the first study of human brain *in vivo* done at such high b-factor and long diffusion times. The persistence of non-monoexponential diffusion decay at very long time is a clear proof that the diffusion in WM **is** compartmental. For the range of diffusion times in our study a significant exchange happens between intra and extra-axonal water but very little exchange is seen for the water inside the glial processes.

From this study three major issues can be addressed: (i) Three components is necessary to describe diffusion decay curves. (ii) Exchange time between all water pools are relatively long but exchange formulation is necessary for proper data fitting at long times. (iii) The slowest of water pool shows isotropic behavior and are most likely water residing inside the glial processes.

[1] When the water in the glial processes end up in the glial soma, it is not exchange between compartments through membranes, but it is still exchange from the perspective of diffusion. The molecules are now not restricted to move in the direction that they were previously moving. The diffusion dimension is increases, which causes and increased ADC.

# Chapter 7

# Conclusions and Outlook

## General Discussion: On the origin of diffusion MRI signal in the brain white matter

In light of the series of preceding measurements, a unifying picture evolves for the diffusion MRI signal in brain WM. The diffusion MR signal can be thought of as arising from five major water pools: (i) myelin water (ii) tubular water (iii) axonal water and (iv) extra-axonal water (v) water in the glial processes.

**Myelin Water** The first water pool is water in the compact myelin. This water fraction resides within the layers of myelin that wrap the axonal lumen and is relatively isolated. The myelin layers contain a relatively high volume fraction of sphingomyelin and cholesterol that make them fairly impermeable to water molecules. Since water is narrowly confined between the two lipid surfaces, it is mostly the hydration layer that are ordered and have greatly reduced dynamics. This fraction of water has slow diffusivity in all directions.

By virtue of the confinement, myelin water is also expected to have a reduced $T_2$ as well as a freezing point depression. The measurements in the excised porcine corpus callosum showed that the 'unfrozen' water fraction had a reduced $T_2$. This in my knowledge is the first direct measurement of myelin water. Due to its short $T_2$, myelin water is only visible at short echo times measurements.

**Tubular Water** Like myelin water, mobility in the second pool is also very 'slow' in all directions. But their $T_2$ times are longer. For measurements at long echo times when myelin water is no longer visible, this water population can be distinguished as remaining slow water pool. We assign this pool to water associated with neuronal microtubules together with some other water trapped in small clefts. Microtubules with a packing density of around $70/\,\mu\mathrm{m}^2$, extend up to hundreds of micrometers in

length and form bundles known as stress fibers [180, 181]. They proliferate all cellular structures, but within axons they are found to be more dense and linearly organized. The water molecules in the vicinity of such nano-scale tubular structure experience a large effective surface to volume ratio. With 13 nm inner diameter almost a third of the water residing inside the microtubular lumen is hydration water in contact with the cylindrical walls. Mobility of this two-dimensionally restricted water inside the microtubules is also hindered by the macromolecules that fill up the lumen. Due to surface interactions, interfacial water is expected to experience a significant drop in its diffusion properties [58]. The water trapped in the inner section of the microtubules have negligible diffusion in the transverse direction so little exchange is expected between them and the water outside. They are also covered with wispy radial outgrowths [251] that trap water as clathrate guest. Introduction of colchincine, an microtubule destroying drug has been found to increase intra-axonal ADC only in the direction parallel to the fibers [77].

A study from Beaulieu and Allen [107] reported that the presence of neurofilaments is not the main determinant to diffusion anisotropy in WM. Filament water has very slow diffusion coefficient in all directions and a small volume fraction, therefore, although the reduction in ADC is not very high, the obstructions still should contribute to some anisotropy of the water inside. Intermediate filaments that are also found to be densely packed in the oligodendritic processes. Our measurement was not able to directly measure diffusion parallel to these processes, but for our observations restrictions due to filaments was not necessary to explain the signal behavior.

The tubules, filaments and also some water trapped in membranes can be together grouped as hydration water. In literature, they have also been put forth as the main reason for slow diffusion in brain [82]. The measurements performed at very long times showed that non-monoexponential decay persisted at even such long times and was relatively stable against perturbations in diffusion times and gradient strength. This showed that these microscopic structures perhaps play a bigger role in the reduction of the ADC but not a major role for non-monoexponential signal behavior.

**Axonal Water** The third water pool is assigned to the water residing inside the axonal lumen. The diffusion in this pool is highly anisotropic - the water molecules can move reasonably fast parallel to the fiber direction, but are restricted radially. In the parallel direction the axonal water is associated with the fast component but in perpendicular direction it is associated with the slow component. If the intra-axonal water is fairly restricted, then its radial ADC is expected to be very small for all feasible diffusion times [252]. Although diffusion is restriced in the perpendicular direction, due to fiber dispersion [229], axonal undulation and other non-linear features in the tissue structure, there will be some residual ADC even when the mean fiber direction is perpendicular to

the measured diffusion direction. As previously mentioned, dispersion of axonal fibers, is expected to result in an elevated radial ADC that is around 6% of the axial ADC.

The myelin forming phospholipid bilayers are expected to have low permeability coefficients across the bilayer and myelin layers are also not known to have aquaporins. It is also known that in WM, radial ADC is increased in the absence of myelin [76, 185, 253].

The myelin layers do not allow free-mixing of intra and extra-axonal water pools so most exchange probably takes place through the node of Ranvier. The node of Ranvier is the primary site for astroglial contact and contains dense clusters of ion channels. However, the nodes are relatively sparse - approximately 1000 nodes are present in a meter long myelinated nerve fiber. They occupy about one thousandth of the axonal surface area [91]. On average the intra-axonal water has to travel long distances before mixing with water in the extra-axonal water which is reflected in long exchange times estimated. A Monte Carlo study that assumed exchange between intra and extra axonal water only through nodes of Ranvier concluded that Kärger equation was sufficient to model such exchange [254]. Their simulations resulted in intra-axonal exchange times that were in the order of thousands of milliseconds [254]. These values are similar to our findings. The inter-nodal distance is known to be closely related to axon caliber [92] and axon diameters in primate CNS are known not to follow a normal distribution [255]. Therefore, a large variability in the exchange time is not too surprising. It is noted that variations in the inter-nodal distance might not be fully grasped with a single exchange parameter.

The periaxonal space filled with extra cellular fluid is still located in the cylindrical region inside the myelin layer. The water residing in the periaxonal space have to traverse a long radial distance before getting out of the myelin encased region. Therefore, for diffusion purpose the water in this space should also be considered in the same pool as intraaxonal water.

**Water in Glial Processes** The water in the glial processes have many similarities to the axonal water. Inside the processes they encounter many cytoskeletal proteins and cell organelles that act as hindrances to their motion. These hindrances are also alligned along the direction of the processes, so that they might contribute to some anisotropy. However, the most prominent features are probably the membranes that do not let the water molecules escape.

But unlike axons, the glial processes are not coherently oriented but more or less randomly distributed. In a geometric sense, a fraction of these glial processes are always parallel to the measured direction and a part is always perpendicular. Thus they contribute both to the fast and the slow pool.

The glial processes are the major contributors to the non-monoexponential signal in the parallel direction. Together with the axons, they also contribute to some of the slow fraction in the perpendicular direction.

**Extra Axonal Water** The last pool comprises of water that is relative free in all directions but still has some anisotropy. Parallel to the fiber direction its value is around 60% of free water (See Fig. 6.8) which is equivalent to a tortuosity of $0.6^2 = 0.36$. Perpendicular to the fiber direction the diffusion coefficient of this pool is about half its value in the parallel direction.. When measured perpendicular to the fibers, this is the only pool that contributes to the fast component and can be easily distinguished from other pools.

The extra-axonal space houses neuroglia - a vast multitude of cells including microglia, and somas of astrocytes and oligodendrocytes. The extracellular space which is very small is also part of this compartment. Extra-axonal volume is mostly occupied by astrocytes and oligodendrocytes. The water inside these glial have enough room within them, so that they are likely to be indistinguishable from the first pool of water. Perhaps, at very long time, the water in these cell bodies face some restrictions from the cell wall which might be the cause of decreasing ADC with increasing diffusion time. However, such predictions are probably still beyond our present measurement techniques.

### 7.0.5 Volume Fractions

The volume of different cellular substructures of the WM have been studied in detail. However, the volume of water contained in these structures lacks in-depth studies. Water density is thought to be uniform throughout the brain, occupying the all the void spaces between the macromolecules.

The volume fraction of the myelin water estimated from porcine corpus callosum was about 19% from $T_2$ measurements and about 14% from freezing. The lipid bilayers are about 5 nm thick and the periodicity of myelin is about 16.6 nm. Water in found in two layer with total thickness of 6.5 nm. This makes the water content of myelin to be around 13/32. The myelin water fraction is also similar to what has been previously reported in literature [170].

The volume of myelin hints on the volume of the axons. The ratio between diameter of axon and the fiber diameter including both axon and myelin layers is called a g-ratio. The g-ratio is considered imported for saltatory conduction in the axons. In brain WM the g-ratio is around 0.78 [256]. Ignoring the periaxonal space which has very small volume compared to the axonal water, the ratio between axonal volume and the myelin volume is then $0.78^2 : 1 - 0.78^2$. Since the axonal lumen is cytoplasmic, due to macromolecular crowding around 75% volume is occupied by water [80]. This means the ratio between axonal water and myelin water is almost about 5:2. Considering the myelin

water fraction to be around 16.5% ( mean of 14 and 19), we get axonal water fraction to be about 40%. Rest of the water is in the extra-axonal space and glial processes is then about 54%.

The relative volume fraction of the extra-axonal water as estimated from the perpendicular direction is slightly more than 50%(Fig. 6.8 and Fig. 6.23). This is congruent with anatomical findings: in optic nerve, myelin, axons and extra-axonal space and glial processes have been reported to be occupy 25%, 32% , 23% and 20% of the WM volume [83]. Since myelin water has negligible signal contribution at long echo times, ignoring myelin water, intra and extra axonal water should contribute to the total signal at a 60/40 volume ratio. Additionally, the glial processes should contribute a similar volume fraction of the signal like the glial soma. In a biexponential fit, the glial process contribute both to the fast and the slow component.

### 7.0.6 Problems with exponential decays

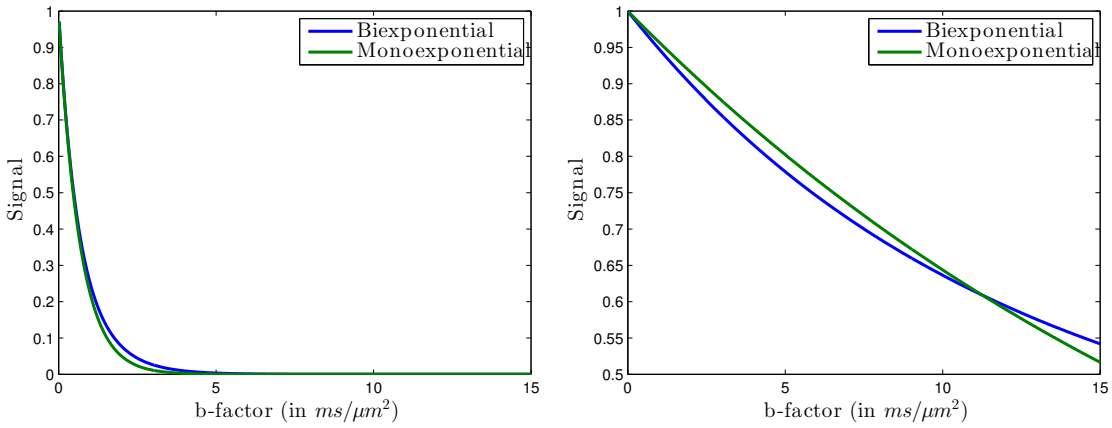



FIGURE 7.1: Proximity between biexponential and monoexponential decay curves. On the left, the biexponential decay arises from two components $D_f = 2.0$ and $D_s = 1.0$ and the monoexponential decay is from a single component with $D = 1.5$. Right: the biexponential decay arises from two components $D_f = 0.1$ and $D_s = 0.01$ and the monoexponential decay is from a single component with $D = 0.045$.

The difficulty of distinguishing different water pools is largely due to the exponential nature of diffusion decay curves. It is well known that fitting data with sum of exponential decays is ill-conditioned problem. Since the data decays exponentially, obtaining meaningful results require sampling of the decay for few orders of magnitude. For example, Fig. 7.1(left) shows two diffusion decay curves, one arising from two diffusion coefficients ( $D_f = 2.0$ and $D_s = 1.0$ ) with equal volume fractions and the other with a single coefficient that is the mean of the two ( $ADC = 1.0$). Even in the absence of noise, the maximum deviation between the two curves is never more than 0.03. Similarly, Fig. 7.1(right) shows the other extreme case with $D_f = 0.1$ and $D_s = 0.01$. A

monoexponential fit with ADC = 0.045 fits the curve fairly well: difference between the two curves is also never more than 0.03. In both cases, differentiation is further complicated with the addition of a third component.

A useful measure arising from this work is the possibility of distinguishing the glial processes, without necessarily treating the data as Gaussian mixture. Since the diffusion along the processes can be considered to be fairly constant, for clinical measurements with a fixed diffusion time, a measure of the volume fraction of the glial processes can be easily obtained from a process similar to that described in the last experimental section.

### 7.0.7 Are we really measuring diffusion compartments?

The answer is yes. There are many tiny yet complex structures in the WM. Diffusion is sensitive to tiniest differences in structures, and yet what we observe is an ensemble average behaviour from all the spins that are in our region of interest.

Compartments are simply spatial heterogeneities that are too small to be individually detected but big enough for diffusing particles to not be able to sample all of these spaces in the given time. The diffusion decay curves in the WM are rather robust and do not show sharp changes with increased time, gradient strength or temperature within the limits of our measurement.

An exception was the application of very strong gradients perpendicular to the fiber orientation in *ex vivo* measurement. In that case, a sharp unnatural drop in signal was seen with increasing gradients. At first this was exciting and quite some effort was invested to see if they were truly diffusion-diffraction pattern observed in q-space imaging. However, the signal drops were very much affected by TE and the gradient peak strength.

## 7.1 Implication of this work

The aim of this work was to better understand diffusion process in brain WM and how that could relate with the measured signal in diffusion NMR measurements. The exponential nature of the decay curves makes it very hard problem to properly quantify measures of interest. But, at the same time it also allows modeling it with many downward sloping curves. This has allowed for many models to be available.

Since the physics behind the observed decays is not so obvious, models describing unlikely phenomenon are also likely to explain the signal - as long as they satisfy the first condition of being monotonically downward sloping. Technically, even in complex systems, the time evolution of diffusion coefficient itself is enough to characterize diffusion process, however, selection of proper measurement parameters are often hard to estimate. This makes comparisons between models very difficult. Experiments performed

during the course of this work provide additional understanding in the diffusive process itself. Modeling has been intentionally avoided.

The main finding stands as the contribution of the glial processes in the non-monoexponential behavior of the diffusion signal.

The inclusion of isotropically distributed glial process should greatly simplify the modeling process. For example, the three dimensional diffusion propagator can be described with just three components: axonal water, extra-axonal water and the water in the glial processes. The axonal water is just one parameter defined along the fiber and could be zero for perpendicular direction. The extra-axonal water is a tensor, with the primary eigenvector parallel to the axonal direction. The glial-process water is a isotropically distributed signal which also just needs a single parameter to be defined. Such modeling is indeed less complicated than current models that use multiple Gaussians to explain the signal decay [148]. Above all glial processes are physical structures and their relevance in the signal behavior does not require introduction of any new concept[1].

The finding that the glial processes are major contributors to non-monoexponential behavior also affects models that try to describe the diffusion decay data. One such model uses kurtosis in parallel and perpendicular direction to measure diffusion in the intra- and extra-axonal space [203]. With two simple measurements, this modeling procedure promises to provide useful information on the tissue micro-structure and pathology. Inclusion of the signal contribution from the glial processes should be further helpful to make such modeling more specific to the tissue properties. One particular example is the kurtosis in the parallel direction. The kurtosis in the parallel direction can be easily explained by a the signal contribution due to the glial process. However, in Ref. [203], it has been explained by a difference in the ADC between intra- and extra-axonal water. To be consistent, the intra-axonal water had to be considered almost half as slow as the extra-axonal water. Considering WM tissue architecture, it is very difficult to believe that axonal-water that face no direct restriction in the axonal direction could be slower than the water residing in the extra-axonal space where they encounter often encounter membranes in the parallel directions as well.

The work presented here should also have an impacts on the models that already exist in literature and to some that have been used in clinics. The first is the measurement of axonal diameters [146, 257]. The measurements at long times show that measurements perpendicular to the WM fiber direction look very similar between diffusion times of 75 ms and 150 ms and even as long as 1500 ms. The signal is almost always a mixture of multiple compartments with freely moving compartments contributing more to the signal decay than restricted ones. Perhaps it is not a good idea to depend on the non-monoexponential diffusion decay curves to estimate axonal diameters.

[1] If the processes themselves follow a random walk then the random walk on those random walk has its own time dependent behavior

The DTI measurements at long diffusion times also showed that the diffusion in brain WM was probably not fractal in nature. It was also noted that fiber dispersion could not be the cause of non-monoexponential diffusion decay in the parallel direction. The body of this work also dealt with the issue of hydration water being responsible for the slow diffusion decay. Presence of a persistent non-monoexponential signal decay at very long diffusion times was taken as a suggestion that these are not the major contributors to the non-monoexponential diffusion decay that is seen in the literature.

The temperature dependence of the diffusion compartments in the brain WM was performed for the first time. Although the activation energy of the slow pool might have mixed information from different pools. Activation energies of both the pools and thereby the activation of the ADC in brain WM is less than that of pure water. Temperature dependence of diffusion is a well established physical phenomenon, but its use in clinic is not heard of. Further research in this direction could find ways to characterize tissue pathology in biopsies.

During the course of this work two main findings that are not directly related to the main body of this work were (i) verification of myelin water as the short $T_2$ technique using freezing technique, and (ii) verification that the physiological noise in the brain is mostly due to CSF contribution.

Towards efficient measurement and analysis, I showed that once the correct model is known, using the model together with the Rician noise parameter is a better and more efficient way to fit the data.

## 7.2 Further verifications for the contribution from glial processes?

The conclusion presented in this thesis is already based on measurements that could be performed with basic diffusion measurements. Thereby, further verifications of this phenomenon would require measurements that combine two measurements into one.

A simple method is to use isotropically weighted diffusion gradients in a single measurement. If the isotropically weighted signal approaches zero at relatively small b-factors, it shows that there is no diffusion pool that is completely restricted. This would be a direct proof of a lack of isotropically restricted component but not a proof for the glial processes to be a major cause of non-monoexponential signal decay.

While isotropic weighting applies diffusion gradient in all directions, one could do a similar measurement with double PFG gradients. Diffusion weighting in parallel direction can suppress signal from extra- and intra-axonal water leaving only water from those glial processes that are more or less perpendicular to the gradient direction. Application of another set of gradients should then show anisotropy but in the direction

perpendicular to the main fiber direction. The tensor would also be oblate - like a disc, higher diffusivity perpendicular to the fiber and lower parallel to it. The initial gradient directions can also be at an angle to the fiber direction. As in our last measurement, the signal can still be suppressed with appropriate diffusion weighting. It should be noted that such measurements require high SNR.

In my opinion, the conclusions arrived in this thesis and the experiments performed have been fairly minimalist. The power in these measurement techniques come from their simplicity that require minimal data manipulation, processing and modelling. By the way of measurements, this thesis shows that diffusion MR at high field *in vivo* is possible and perhaps stands to further benefit from better understanding of the tissue microstructure that are relevant to the measurement.

# Bibliography


[1] Brown R. A brief account of microscopical observations made in the months of june, july, and august 1827 on the particles contained in the polen of plants;and on the general existence of active molecules in organic and inorganic bodies. Philosophical Magazine 1828;4:161.

[2] Einstein A. Investigations on the theory of brownian motion. dover, new york, 1956. (collection of papers translated from the german). 1956.

[3] de Moivre A. Miscellanea analytica de seriebus et quadraturis. J. Tonson & J. Watts 1730.

[4] Diaconis P, Freedman D. An elementary proof of stirling's formula. American Mathematical Monthly 1986;pages 123–125.

[5] Fick D. On liquid diffusion. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science 1855;10:30–39.

[6] Dvinskikh S V, Furo I. Anisotropic self-diffusion in the nematic phase of a thermotropic liquid crystal by h-spin-echo nuclear magnetic resonance. The Journal of Chemical Physics 2001;115:1946.

[7] Onsager L. Electric moments of molecules in liquids. Journal of the American Chemical Society 1936;58:1486–1493.

[8] Bowman F. Introduction to Bessel functions. Courier Dover Publications 1958.

[9] Mitra P P, Sen P N, Schwartz L M, Le Doussal P. Diffusion propagator as a probe of the structure of porous media. Physical review letters 1992;68:3555–3558.

[10] Mitra P P, Sen P N, Schwartz L M. Short-time behavior of the diffusion coefficient as a geometrical probe of porous media. Physical Review B 1993;47:8565.

[11] Latour L L, Mitra P P, Kleinberg R L, Sotak C H. Time-dependent diffusion coefficient of fluids in porous media as a probe of surface-to-volume ratio. Journal of Magnetic Resonance, Series A 1993;101:342–346.

[12] Latour L L, Svoboda K, Mitra P P, Sotak C H. Time-dependent diffusion of water in a biological model system. Proceedings of the National Academy of Sciences 1994;91:1229–1233.

[13] de Swiet T M, Sen P N. Time dependent diffusion coefficient in a disordered medium. The Journal of chemical physics 1996;104:206–209.

[14] Sen P N. Time-dependent diffusion coefficient as a probe of geometry. Concepts in Magnetic Resonance Part A 2004;23:1–21.

[15] Novikov D S, Jensen J H, Helpern J A, Fieremans E. Revealing mesoscopic structural universality with diffusion. Proceedings of the National Academy of Sciences 2014;111:5088–5093.

[16] Woessner D E. NMR spin-echo self-diffusion measurements on fluids undergoing restricted diffusion. The Journal of Physical Chemistry 1963;67:1365–1367.

[17] Le Bihan D, Mangin J F, Poupon C, Clark C A, Pappata S, Molko N, Chabriat H. Diffusion tensor imaging: concepts and applications. Journal of magnetic resonance imaging 2001;13:534–546.

[18] Soppimath K S, Aminabhavi T M, Kulkarni A R, Rudzinski W E. Biodegradable polymeric nanoparticles as drug delivery devices. Journal of controlled release 2001;70:1–20.

[19] Levitt M H. Spin dynamics: basics of nuclear magnetic resonance. John Wiley & Sons 2008.

[20] Rabi I. Space quantization in a gyrating magnetic field. Physical Review 1937;51:652.

[21] Bloch F, Hansen W, Packard M. Nuclear induction. Physical review 1946;70:460–474.

[22] Hahn E. Spin echoes. Physical Review 1950;80:580.

[23] Carr H, Purcell E. Effects of diffusion on free precession in nuclear magnetic resonance experiments. Physical Review 1954;94:630.

[24] Torrey H. Bloch equations with diffusion terms. Physical Review 1956;104:563.

[25] Stejskal E, Tanner J. Spin diffusion measurements: Spin echoes in the presence of a time-dependent field gradient. The journal of chemical physics 1965;42:288.

[26] Rabi I, Ramsey N, Schwinger J. Use of rotating coordinates in magnetic resonance problems. Reviews of Modern Physics 1954;26:167.

[27] Garroway A N, Grannell P K, Mansfield P. Image formation in NMR by a selective irradiative process. Journal of Physics C: Solid State Physics 1974;7:L457.

[28] Hoult D. The solution of the bloch equations in the presence of a varying b1 field—an approach to selective pulse analysis. Journal of Magnetic Resonance (1969) 1979;35:69–86.

[29] Shinnar M. Reduced power selective excitation radio frequency pulses. Magnetic resonance in medicine 1994;32:658–660.

[30] Le Roux P. Exact synthesis of radio frequency waveforms. In: *Proc. 7th SMRM*;volume 1049 1988.

[31] Bandettini P A, Schmitt F, Stehling M K, Turner R. Echo-planar imaging: theory, technique, and application. Springer 1998.

[32] Woessner D E. Self-diffusion measurements in liquids by the spin-echo technique. Review of Scientific Instruments 1960;31:1146–1146.

[33] Winkler H, Gutsze A. On two-component NMR relaxation. Advances in Molecular Relaxation and Interaction Processes 1981;21:159–179.

[34] Karger J. NMR self-diffusion studies in heterogeneous systems. Advances in Colloid and Interface Science 1985;23:129–148.

[35] Jorg Karger, Harry Pfeifer, Wilfried Heink. Principles and applications of self-diffusion measurements by nuclear magnetic resonance. In: John S Waugh, editor. Advances in Magnetic Resonance Academic Press, Inc. 1988;pages 1–89.

[36] Callaghan P T. Pulsed-gradient spin-echo NMR for planar, cylindrical, and spherical pores under conditions of wall relaxation. Journal of magnetic resonance, Series A 1995;113:53–59.

[37] Cory D. Measurement of translational displacement probabilities by NMR: an indicator of compartmentation. Magnetic resonance in medicine 1990;14:435–444.

[38] Callaghan P T, Coy A, MacGowan D, Packer K J, Zelaya F O. Diffraction-like effects in NMR diffusion studies of fluids in porous solids. 1991.

[39] Neuman C. Spin echo of spins diffusing in a bounded medium. The Journal of Chemical Physics 1974;60:4508.

[40] Robertson B. Spin-echo decay of spins diffusing in a bounded region. Physical Review 1966;151:273.

[41] Wayne R C, Cotts R M. Nuclear-magnetic-resonance study of self-diffusion in a bounded medium. Physical Review 1966;151:264.

[42] de Swiet T M, Sen P N. Decay of nuclear magnetization by bounded diffusion in a constant field gradient. The Journal of chemical physics 1994;100:5597–5604.

[43] Hurlimann M, Helmer K, Deswiet T, Sen P. Spin echoes in a constant gradient and in the presence of simple restriction. Journal of Magnetic Resonance, Series A 1995;113:260–264.

[44] Jensen J H, Helpern J A, Ramani A, Lu H, Kaczynski K. Diffusional kurtosis imaging: The quantification of non-gaussian water diffusion by means of magnetic resonance imaging. Magnetic Resonance in Medicine 2005;53:1432–1440.

[45] Kiselev V G, Ilyasov K A. Is the "biexponential diffusion" biexponential? Magnetic Resonance in Medicine 2007;57:464–469.

[46] Frohlich A F, Ostergaard L, Kiselev V G. Effect of impermeable boundaries on diffusion-attenuated MR signal. Journal of Magnetic Resonance 2006;179:223–233.

[47] Brodal P. The central nervous system: structure and function. Oxford University Press 2004.

[48] Nixon R. Cytoskeleton of the nervous system;volume 3. Springer Verlag 2011.

[49] OBrien J S, Sampson E L. Lipid composition of the normal human brain: gray matter, white matter, and myelin. Journal of lipid research 1965;6:537–544.

[50] Lande M B, Donovan J M, Zeidel M L. The relationship between membrane fluidity and permeabilities to water, solutes, ammonia, and protons. The Journal of general physiology 1995;106:67–84.

[51] Hishida M, Tanaka K. Long-range hydration effect of lipid membrane studied by terahertz time-domain spectroscopy. Physical review letters 2011;106:158102.

[52] Cheng J, Pautot S, Weitz D, Xie X. Ordering of water molecules between phospholipid bilayers visualized by coherent anti-stokes raman scattering microscopy. Proceedings of the National Academy of Sciences 2003;100:9826.

[53] Bhide S Y, Berkowitz M L. Structure and dynamics of water at the interface with phospholipid bilayers. The Journal of chemical physics 2005;123:224702.

[54] Englund E, Brun A, Alling C. White matter changes in dementia of Alzheimer's type biochemical and neuropathological correlates. Brain 1988;111:1425–1439.

[55] Alberts B, et al. Molecular biology of the cell in cell 4th;2002.

[56] Gopalakrishnan G, Awasthi A, Belkaid W, De Faria O, Liazoghli D, Colman D R, Dhaunchak A S. Lipidome and proteome map of myelin membranes. Journal of neuroscience research 2013;91:321–334.

[57] Goertz M P, Houston J, Zhu X Y. Hydrophilicity and the viscosity of interfacial water. Langmuir 2007;23:5491–5497.

[58] Settles M, Doster W. Anomalous diffusion of adsorbed water: a neutron scattering study of hydrated myoglobin. Faraday Discussions 1996;103:269–279.

[59] Doster W, Settles M. Protein–water displacement distributions. Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics 2005;1749:173–186.

[60] Zhong D, Pal S K, Zewail A H. Biological water: A critique. Chemical Physics Letters 2011;503:1–11.

[61] Kaplan M R, Meyer Franke A, Lambert S, Bennett V, Duncan I, Levinson S, Barres B. Induction of sodium channel clustering by oligodendrocytes. 1997.

[62] Preston G M, Carroll T P, Guggino W B, Agre P. Appearance of water channels in xenopus oocytes expressing red cell chip28 protein. Science 1992;256:385–387.

[63] Webster R E, Henderson D, Osborn M, Weber K. Three-dimensional electron microscopical visualization of the cytoskeleton of animal cells: immunoferritin identification of actin-and tubulin-containing structures. Proceedings of the National Academy of Sciences 1978;75:5511–5515.

[64] Heuser J E, Kirschner M. Filament organization revealed in platinum replicas of freeze-dried cytoskeletons. The Journal of cell biology 1980;86:212–234.

[65] Luby Phelps K. Cytoarchitecture and physical properties of cytoplasm: volume, viscosity, diffusion, intracellular surface area. International review of cytology 1999;192:189–221.

[66] Santerre R F, Rich A. Actin accumulation in developing chick brain and other tissues. Developmental biology 1976;54:1–12.

[67] Siegel G J, Agranoff B W, Albers R W, Molinoff P B. Basic neurochemistry: molecular, cellular, and medical aspects. Raven Press 1994.

[68] Chou Y H, Skalli O, Goldman R D. Intermediate filaments and cytoplasmic networking: new connections and more functions. Current opinion in cell biology 1997;9:49–53.

[69] Wuerker R B, Kirkpatrick J B. Neuronal microtubules, neurofilaments, and microfilaments. International review of cytology 1972;33:45.

[70] Bryan J, Wilson L. Are cytoplasmic microtubules heteropolymers? Proceedings of the National Academy of Sciences 1971;68:1762–1766.

[71] Hiller G, Weber K. Radioimmunoassay for tubulin: a quantitative comparison of the tubulin content of different established tissue culture cells and tissues. Cell 1978;14:795–804.

[72] Yu W, Baas P W. Changes in microtubule number and length during axon differentiation. The Journal of neuroscience 1994;14:2818–2829.

[73] Filliatreau G, Digiamberardino L. Invivo reassembly of axonal microtubules-a fast unidirectional process;1981.

[74] Karlsson J O, Sjostrand J. The effect of colchicine on the axonal transport of protein in the optic nerve and tract of the rabbit. Brain research 1969;13:617–619.

[75] Bizzozero O A, Pasquini J M, Soto E F. Differential effect of colchicine upon the entry of proteins into myelin and myelin related membranes. Neurochemical research 1982;7:1415–1425.

[76] Beaulieu C, Allen P S. Determinants of anisotropic water diffusion in nerves. Magnetic resonance in medicine 1994;31:394–400.

[77] Kinoshita Y, Ohnishi A, Kohshi K, Yokota A. Apparent diffusion coefficient on rat brain and nerves intoxicated with methylmercury. Environmental research 1999;80:348–354.

[78] Garvalov B K, Zuber B, Bouchet Marquis C, Kudryashev M, Gruska M, Beck M, Leis A, Frischknecht F, Bradke F, Baumeister W, et al. Luminal particles within cellular microtubules. The Journal of cell biology 2006;174:759–765.

[79] Bouchet Marquis C, Zuber B, Glynn A M, Eltsov M, Grabenbauer M, Goldie K N, Thomas D, Frangakis A S, Dubochet J, Chretien D. Visualization of cell microtubules in their native state. Biology of the Cell 2007;99:45–53.

[80] Ellis R J. Macromolecular crowding: obvious but underappreciated. Trends in biochemical sciences 2001;26:597–604.

[81] Schnell S, Turner T. Reaction kinetics in intracellular environments with macromolecular crowding: simulations and rate laws. Progress in biophysics and molecular biology 2004;85:235–260.

[82] Le Bihan D. The'wet mind': water and functional neuroimaging. Physics in medicine and biology 2007;52:R57.

[83] Perge J A, Koch K, Miller R, Sterling P, Balasubramanian V. How the optic nerve allocates space, energy capacity, and information. The Journal of neuroscience 2009;29:7917–7928.

[84] Takahashi M, Hackney D B, Zhang G, Wehrli S L, Wright A C, OBrien W T, Uematsu H, Wehrli F W, Selzer M E. Magnetic resonance microimaging of intraaxonal water diffusion in live excised lamprey spinal cord. Proceedings of the National Academy of Sciences 2002;99:16192–16196.

[85] Robertson J D. Preliminary observations on the ultrastructure of nodes of ranvier. Cell and Tissue Research 1959;50:553–560.

[86] Mierzwa A, Shroff S, Rosenbluth J. Permeability of the paranodal junction of myelinated nerve fibers. The Journal of Neuroscience 2010;30:15962–15968.

[87] Kirschner D, Blaurock A. Organization, phylogenetic variations, and dynamic transitions of myelin. Myelin: biology and chemistry 1992;pages 3–78.

[88] De Felici M, Felici R, Ferrero C, Tartari A, Gambaccini M, Finet S. Structural characterization of the human cerebral myelin sheath by small angle x-ray scattering. Physics in medicine and biology 2008;53:5675.

[89] Jahn O, Tenzer S, Werner H B. Myelin proteomics: molecular anatomy of an insulating sheath. Molecular neurobiology 2009;40:55–72.

[90] Koenig S H. Cholesterol of myelin is the determinant of gray-white contrast in MRI of brain. Magnetic resonance in medicine 1991;20:285–291.

[91] Rasband M N, Trimmer J S. Developmental clustering of ion channels at and near the node of ranvier. Developmental biology 2001;236:5–16.

[92] Hess A, Young J. The nodes of ranvier. Proceedings of the Royal Society of London. Series B, Biological Sciences 1952;pages 301–320.

[93] Torres Platas S G, Hercher C, Davoli M A, Maussion G, Labonte B, Turecki G, Mechawar N. Astrocytic hypertrophy in anterior cingulate white matter of depressed suicides. Neuropsychopharmacology 2011;36:2650–2658.

[94] Wang D D, Bordey A. The astrocyte odyssey. Progress in neurobiology 2008;86:342–367.

[95] Sun D, Lye Barthel M, Masland R H, Jakobs T C. Structural remodeling of fibrous astrocytes after axonal injury. The Journal of Neuroscience 2010;30:14008–14019.

[96] Lunn K F, Baas P W, Duncan I D. Microtubule organization and stability in the oligodendrocyte. The Journal of neuroscience 1997;17:4921–4932.

[97] Callaghan P, Jolley K, Lelievre J. Diffusion of water in the endosperm tissue of wheat grains as studied by pulsed field gradient nuclear magnetic resonance. Biophysical journal 1979;28:133–141.

[98] Haseleu J, Anlauf E, Blaess S, Endl E, Derouiche A. Studying subcellular detail in fixed astrocytes: dissociation of morphologically intact glial cells (dimigs). Frontiers in cellular neuroscience 2013;7.

[99] Jespersen S, Kroenke C, Ostergaard L, Ackerman J, Yablonskiy D. Modeling dendrite density from magnetic resonance diffusion measurements. Neuroimage 2007;34:1473–1486.

[100] Yablonskiy D A, Sukstanskii A L, Leawoods J C, Gierada D S, Bretthorst G L, Lefrak S S, Cooper J D, Conradi M S. Quantitative in vivo assessment of lung microstructure at the alveolar level with hyperpolarized 3he diffusion MRI. Proceedings of the National Academy of Sciences 2002;99:3111–3116.

[101] Basser P J, Mattiello J, LeBihan D. MR diffusion tensor spectroscopy and imaging. Biophysical journal 1994;66:259–267.

[102] Pierpaoli C, Basser P J. Toward a quantitative assessment of diffusion anisotropy. Magnetic resonance in Medicine 1996;36:893–906.

[103] Westin C F, Maier S E, Mamata H, Nabavi A, Jolesz F A, Kikinis R. Processing and visualization for diffusion tensor MRI. Medical image analysis 2002;6:93–108.

[104] Mitra P P, Halperin B I. Effects of finite gradient-pulse widths in pulsed-field-gradient diffusion measurements. Journal of Magnetic Resonance, Series A 1995;113:94–101.

[105] Norris D G. The effects of microscopic tissue parameters on the diffusion weighted magnetic resonance imaging experiment. NMR in Biomedicine 2001;14:77–93.

[106] Wiegell M R, Larsson H B, Wedeen V J. Fiber crossing in human brain depicted with diffusion tensor MR imaging 1. Radiology 2000;217:897–903.

[107] Beaulieu C, Allen P S. Water diffusion in the giant axon of the squid: implications for diffusion-weighted MRI of the nervous system. Magnetic Resonance in Medicine 1994;32:579–583.

[108] Gulani V, Webb A, Duncan I, Lauterbur P. Apparent diffusion tensor measurements in myelin-deficient rat spinal cords. Magnetic Resonance in Medicine 2001;45:191–195.

[109] Wimberger D M, Roberts T P, Barkovich A J, Prayer L M, Moseley M E, Kucharczyk J. Identification of" premyelination" by diffusion-weighted MRI. Journal of computer assisted tomography 1995;19:28–33.

[110] Partridge S C, Mukherjee P, Henry R G, Miller S P, Berman J I, Jin H, Lu Y, Glenn O A, Ferriero D M, Barkovich A J, et al. Diffusion tensor imaging: serial quantitation of white matter tract maturity in premature newborns. Neuroimage 2004;22:1302–1314.

[111] Tuch D S, Wedeen v J, Dale A M, George J S, Belliveau J W. Conductivity mapping of biological tissue using diffusion MRI. Annals of the New York Academy of Sciences 1999;888:314–316.

[112] Le Bihan D, Turner R, Douek P. Is water diffusion restricted in human brain white matter? an echo-planar NMR imaging study. Neuroreport 1993;4:887.

[113] Moonen C T, Pekar J, de Vleeschouwer M H, van Gelderen P, van Zijl P, Despres D. Restricted and anisotropic displacement of water in healthy cat brain and in stroke studied by NMR diffusion imaging. Magnetic resonance in medicine 1991;19:327–332.

[114] Pfeuffer J, Dreher W, Sykova E, Leibfritz D. Water signal attenuation in diffusion-weighted¡ sup¿ 1¡/sup¿ h NMR experiments during cerebral ischemia: influence of intracellular restrictions, extracellular tortuosity, and exchange. Magnetic resonance imaging 1998;16:1023–1032.

[115] Pfeuffer J, Provencher S, Gruetter R. Water diffusion in rat brain in vivo as detected at very largeb-values is multicompartmental. Magnetic Resonance Materials in Physics, Biology and Medicine 1999;8:98–108.

[116] Meier C, Dreher W, Leibfritz D. Diffusion in compartmental systems. ii. diffusion-weighted measurements of rat brain tissue in vivo and postmortem at very large b-values. Magnetic resonance in medicine 2003;50:510–514.

[117] Assaf Y, Cohen Y. Detection of different water populations in brain tissue using¡ sup¿ 2¡/sup¿ h single-and double-quantum-filtered diffusion NMR spectroscopy. Journal of Magnetic Resonance, Series B 1996;112:151–159.

[118] Niendorf T, Dijkhuizen R M, Norris D G, van Lookeren Campagne M, Nicolay K. Biexponential diffusion attenuation in various states of brain tissue: Implications for diffusion-weighted imaging. Magnetic Resonance in Medicine 1996;36:847–857.

[119] Clark C, Le Bihan D. Water diffusion compartmentation and anisotropy at high b-values in the human brain. Magnetic Resonance in Medicine 2000;44:852–859.

[120] Mulkern R, Gudbjartsson H, Westin C, Zengingonul H, Gartner W, Guttmann C, Robertson R, Kyriakos W, Schwartz R, Holtzman D, et al. Multi-component apparent diffusion coefficients in human brain$^2$. NMR Biomed 1999;12:51–62.

[121] Mulkern R, Haker S, Maier S. On high b diffusion imaging in the human brain: ruminations and experimental insights. Magnetic resonance imaging 2009;27:1151–1162.

[122] Assaf Y, Cohen Y. Assignment of the water slow-diffusing component in the central nervous system using q-space diffusion MRS: Implications for fiber tract imaging. Magnetic resonance in medicine 2000;43:191–199.

[123] Alexander A, Hasan K, Lazar M, Tsuruda J, Parker D. Analysis of partial volume effects in diffusion-tensor MRI. Magnetic Resonance in Medicine 2001;45:770–780.

[124] Stanisz G, Wright G, Henkelman R, Szafer A. An analytical model of restricted diffusion in bovine optic nerve. Magnetic Resonance in Medicine 1997;37:103–111.

[125] Mulkern R, Zengingonul H, Robertson R, Bogner P, Zou K, Gudbjartsson H, Guttmann C, Holtzman D, Kyriakos W, Jolesz F, et al. Multi-component apparent diffusion coefficients in human brain: relationship to spin-lattice relaxation. Magnetic Resonance in Medicine 2000;44:292–300.

[126] Mulkern R V, Vajapeyam S, Haker S J, Maier S E. Magnetization transfer studies of the fast and slow tissue water diffusion components in the human brain. NMR in Biomedicine 2004;18:186–194.

[127] Inglis B, Bossart E, Buckley D, Wirth III E, Mareci T. Visualization of neural tissue water compartments using biexponential diffusion tensor MRI. Magnetic resonance in medicine 2001;45:580–587.

[128] Clark C, Hedehus M, Moseley M. In vivo mapping of the fast and slow diffusion tensors in human brain. Magnetic resonance in Medicine 2002;47:623–628.

[129] Maier S, Vajapeyam S, Mamata H, Westin C, Jolesz F, Mulkern R. Biexponential diffusion tensor analysis of human brain diffusion data. Magnetic resonance in medicine 2004;51:321–330.

[130] Assaf Y, Ben Bashat D, Chapman J, Peled S, Biton I, Kafri M, Segev Y, Hendler T, Korczyn A, Graif M, et al. High b-value q-space analyzed diffusion-weighted MRI: Application to multiple sclerosis. Magnetic resonance in medicine 2002;47:115–126.

[131] Assaf Y, Mayzel Oreg O, Gigi A, Ben Bashat D, Mordohovitch M, Verchovsky R, Reider Groswasser I, Hendler T, Graif M, Cohen Y, et al. High b-value q-space

analyzed diffusion MRI in vascular dementia: A preliminary study. Journal of the neurological sciences 2002;203:235–239.

[132] Maier S, Bogner P, Bajzik G, Mamata H, Mamata Y, Repa I, Jolesz F, Mulkern R. Normal brain and brain tumor: Multicomponent apparent diffusion coefficient line scan imaging1. Radiology 2001;219:842–849.

[133] Assaf Y, Freidlin R, Rohde G, Basser P. New modeling and experimental framework to characterize hindered and restricted water diffusion in brain white matter. Magnetic Resonance in Medicine 2004;52:965–978.

[134] Vangelderen P, DesPres D, Vanzijl P, Moonen C. Evaluation of restricted diffusion in cylinders. phosphocreatine in rabbit leg muscle. Journal of Magnetic Resonance, Series B 1994;103:255–260.

[135] Assaf Y, Basser P J. Composite hindered and restricted model of diffusion (charmed) MR imaging of the human brain. Neuroimage 2005;27:48–58.

[136] Assaf Y, Cohen Y, et al. In vivo and in vitro bi-exponential diffusion of n-acetyl aspartate(naa) in rat brain: a potential structural probe? NMR in Biomedicine 1998;11:67–74.

[137] Nezamzadeh M. Diffusion time dependence of magnetic resonance diffusion signal decays: an investigation of water exchange in human brain in vivo. Magnetic Resonance Materials in Physics, Biology and Medicine 2012;pages 1–12.

[138] Nossin Manor R, Duvdevani R, Cohen Y. Effect of experimental parameters on high b-value q-space MR images of excised rat spinal cord. Magnetic resonance in medicine 2005;54:96–104.

[139] Bar Shir A, Cohen Y. High b-value q-space diffusion MRS of nerves: structural information and comparison with histological evidence. NMR in biomedicine 2007;21:165–174.

[140] Cohen Y, Assaf Y. High b-value q-space analyzed diffusion-weighted MRS and MRI in neuronal tissues–a technical review. NMR in Biomedicine 2002;15:516–542.

[141] Ong H, Wright A, Wehrli S, Souza A, Schwartz E, Hwang S, Wehrli F. Indirect measurement of regional axon diameter in excised mouse spinal cord with q-space imaging: Simulation and experimental studies. NeuroImage 2008;40:1619–1632.

[142] Bar Shir A, Avram L, Ozarslan E, Basser P, Cohen Y. The effect of the diffusion time and pulse gradient duration ratio on the diffraction pattern and the structural

information estimated from q-space diffusion MR: experiments and simulations. Journal of Magnetic Resonance 2008;194:230–236.

[143] Amnon Bar Shir, Yoram Cohen. The effect of the rotational angle on MR diffusion indices in nerves: Is the rms displacement of the slow-diffusing component a good measure of fiber orientation? Journal of Magnetic Resonance January 2008;190:33–42.

[144] Barazany D, Basser P J, Assaf Y. In vivo measurement of axon diameter distribution in the corpus callosum of rat brain. Brain 2009;132:1210–1220.

[145] Ong H, Wehrli F. Quantifying axon diameter and intra-cellular volume fraction in excised mouse spinal cord with q-space imaging. NeuroImage July 2010;51:1360–1366.

[146] Assaf Y, Blumenfeld T, Levin G, Yovel Y, Basser P. Axcaliber - a method to measure the axon diameter distribution and density in neuronal tissues. In: *Proc Intl Soc Magn Reson Med*;volume 14;page 637. Citeseer 2006.

[147] Assaf Y, Blumenfeld Katzir T, Yovel Y, Basser P J. Axcaliber: a method for measuring axon diameter distribution from diffusion MRI. Magnetic Resonance in Medicine 2008;59:1347–1354.

[148] Panagiotaki E, Schneider T, Siow B, Hall M, Lythgoe M, Alexander D. Compartment models of the diffusion MR signal in brain white matter: A taxonomy and comparison. NeuroImage 2011.

[149] Lu H, Jensen J H, Ramani A, Helpern J A. Three-dimensional characterization of non-gaussian water diffusion in humans using diffusion kurtosis imaging. NMR in Biomedicine 2006;19:236–247.

[150] Schultz T, Fuster A, Ghosh A, Deriche R, Florack L, Lek Heng L, et al. Higher-order tensors in diffusion imaging. In: *Visualization and Processing of Tensors and Higher Order Descriptors for Multi-Valued Data. Dagstuhl Seminar 2011* 2013.

[151] Pollack G H, Chin W C. Phase Transitions in Cell Biology. Springer 2008.

[152] Pollack G H, Cameron I L, Wheatley D N. Water and the Cell. Springer 2007.

[153] Naguib N, Ye H, Gogotsi Y, Yazicioglu A G, Megaridis C M, Yoshimura M. Observation of water confined in nanometer channels of closed carbon nanotubes. Nano Letters 2004;4:2237–2243.

[154] Beaulieu C, Allen P. Determinants of anisotropic water diffusion in nerves. Magnetic resonance in medicine 1994;31:394–400.

[155] Henkelman R, Stanisz G, Kim J, Bronskill M. Anisotropy of NMR properties of tissues. Magnetic resonance in medicine 1994;32:592–601.

[156] Alexander D. A general framework for experiment design in diffusion MRI and its application in measuring direct tissue-microstructure features. Magnetic Resonance in Medicine 2008;60:439–448.

[157] Holz M, Heil S, Sacco A. Temperature-dependent self-diffusion coefficients of water and six selected molecular liquids for calibration in accurate 1H NMR PFG measurements. Phys. Chem. Chem. Phys. 2000;2:4740–4742.

[158] Self-diffusion and structure of liquid water. iii. measurement of the self-diffusion of liquid water with [2h.

[159] Gillen K, Douglass D, Hoch M. Self-Diffusion in Liquid Water to -31 °C. The Journal of Chemical Physics 1972;57:5117.

[160] Morishige K, Kawano K. Freezing and melting of water in a single cylindrical pore: The pore-size dependence of freezing and melting behavior. The Journal of Chemical Physics 1999;110:4867.

[161] Ashworth E, Abeles F. Freezing behavior of water in small pores and the possible role in the freezing of plant tissues. Plant Physiology 1984;76:201–204.

[162] Petrov O, Furo I. NMR cryoporometry: Principles, applications and potential. Prog. Nucl. Magn. Reson. Spectrosc 2009;54:97–122.

[163] Karger J, Bar N, Heink W, Pfeifer H, Seiffert G. On the use of pulsed field gradients in a high-field NMR spectrometer to study restricted diffusion in zeolites. Zeitschrift für Naturforschung. A 1995;50:186–190.

[164] Galvosas P, Stallmach F, Seiffert G, Karger J, Kaess U, Majer G. Generation and application of ultra-high-intensity magnetic field gradient pulses for NMR spectroscopy. Journal of Magnetic Resonance 2001;151:260–268.

[165] Meiboom S, Gill D. Modified spin-echo method for measuring nuclear relaxation times. Review of Scientific Instruments 1958;29:688–691.

[166] Lawson C, Hanson R. Solving least squares problems;volume 1. Prentice-Hall Series in Automatic Computation, Englewood Cliffs: Prentice-Hall, 1974 1974.

[167] Provencher S. A constrained regularization method for inverting data represented by linear algebraic or integral equations. Computer Physics Communications 1982;27:213–227.

[168] Blachere J, Young J. The freezing point of water in porous glass. Journal of the American Ceramic Society 1972;55:306–308.

[169] Whittall K, Mackay A, Graeb D, Nugent R, Li D, Paty D. In vivo measurement of T2 distributions and water contents in normal human brain. Magnetic resonance in medicine 1997;37:34–43.

[170] MacKay A, Laule C, Vavasour I, Bjarnason T, Kolind S, Madler B. Insights into brain microstructure from the T2 distribution. Magnetic resonance imaging 2006;24:515–525.

[171] Fung B, McGaughy T. Study of spin-lattice and spin-spin relaxation times of 1H, 2H, and 17O in muscular water. Biophysical journal 1979;28:293–303.

[172] Labadie C, Lee J H, Rooney W D, Jarchow S, Aubert Frecon M, Springer C S, Moller H E. Myelin water mapping by spatially regularized longitudinal relaxographic imaging at high magnetic fields. Magnetic Resonance in Medicine 2014;71:375–387.

[173] Bellissent Funel M. Hydration processes in biology: Theoretical and experimental approaches;volume 305. Ios Pr Inc 1999.

[174] Escanye J, Canet D, Robert J. Nuclear magnetic relaxation studies of water in frozen biological tissues. cross-relaxation effects between protein and bound water protons. Journal of Magnetic Resonance (1969) 1984;58:118–131.

[175] Smith J, Cappa C, Wilson K, Messer B, Cohen R, Saykally R. Energetics of hydrogen bond network rearrangements in liquid water. Science 2004;306:851–853.

[176] Ohmine I, Saito S. Water dynamics: Fluctuation, relaxation, and chemical reactions in hydrogen bond network rearrangement. Accounts of chemical research 1999;32:741–749.

[177] Peled S, Cory D G, Raymond S A, Kirschner D A, Jolesz F A. Water diffusion, T2, and compartmentation in frog sciatic nerve. Magnetic resonance in medicine: official journal of the Society of Magnetic Resonance in Medicine/Society of Magnetic Resonance in Medicine 1999;42:911.

[178] Milne M, Conradi M. Multi-exponential signal decay from diffusion in a single compartment. Journal of Magnetic Resonance 2009;197:87–90.

[179] Sukstanskii A, Ackerman J, Yablonskiy D. Effects of barrier-induced nuclear spin magnetization inhomogeneities on diffusion-attenuated MR signal. Magnetic resonance in medicine 2003;50:735–742.

[180] Wade R. On and around microtubules: an overview. Molecular biotechnology 2009;43:177–191.

[181] Bartlett W, Banker G. An electron microscopic study of the development of axons and dendrites by hippocampal neurons in culture. ii. synaptic relationships. The Journal of neuroscience 1984;4:1954–1965.

[182] Behrens T, Johansen Berg H, Woolrich M, Smith S, Wheeler Kingshott C, Boulby P, Barker G, Sillery E, Sheehan K, Ciccarelli O, et al. Non-invasive mapping of connections between human thalamus and cortex using diffusion imaging. Nature neuroscience 2003;6:750–757.

[183] Bashat D B, Sira L B, Graif M, Pianka P, Hendler T, Cohen Y, Assaf Y. Normal white matter development from infancy to adulthood: comparing diffusion tensor and high b-value diffusion weighted MR images. Journal of magnetic resonance imaging 2005;21:503–511.

[184] Mulkern R, Vajapeyam S, Robertson R, Caruso P, Rivkin M, Maier S. Biexponential apparent diffusion coefficient parametrization in adult vs newborn brain. Magnetic resonance imaging 2001;19:659–668.

[185] Mukherjee P, McKinstry R C. Diffusion tensor imaging and tractography of human brain development. Neuroimaging Clinics of North America 2006;16:19–43.

[186] Kono K, Inoue Y, Nakayama K, Shakudo M, Morino M, Ohata K, Wakasa K, Yamada R. The role of diffusion-weighted imaging in patients with brain tumors. American journal of neuroradiology 2001;22:1081–1088.

[187] Yoshiura T, Mihara F, Tanaka A, Ogomori K, Ohyagi Y, Taniwaki T, Yamada T, Yamasaki T, Ichimiya A, Kinukawa N, et al. High b-value diffusion-weighted imaging is more sensitive to white matter degeneration in Alzheimer's disease. Neuroimage 2003;20:413–419.

[188] Maier S E, Sun Y, Mulkern R V. Diffusion imaging of brain tumors. NMR in biomedicine 2010;23:849–864.

[189] Turner R, Le Bihan D, Maier J, Vavrek R, Hedges L, Pekar J, et al. Echo-planar imaging of intravoxel incoherent motion. Radiology 1990;177:407–414.

[190] Maier S, Mulkern R. Biexponential analysis of diffusion-related signal decay in normal human cortical and deep gray matter. Magnetic resonance imaging 2008;26:897–904.

[191] Reese T, Heid O, Weisskoff R, Wedeen V. Reduction of eddy-current-induced distortion in diffusion MRI using a twice-refocused spin echo. Magnetic Resonance in Medicine 2003;49:177–182.

[192] Feinberg D A, Jakab P D. Tissue perfusion in humans studied by fourier velocity distribution, line scan, and echo-planar imaging. Magnetic resonance in medicine 1990;16:280–293.

[193] Seehafer J, Kalthoff D, Farr T, Wiedermann D, Hoehn M. No increase of the blood oxygenation level-dependent functional magnetic resonance imaging signal with higher field strength: implications for brain activation studies. The Journal of Neuroscience 2010;30:5234–5241.

[194] Wansapura J, Holland S, Dunn R, Ball Jr W. NMR relaxation times in the human brain at 3.0 Tesla. Journal of magnetic resonance imaging 1999;9:531–538.

[195] Tanner J. Use of the stimulated echo in NMR diffusion studies. The Journal of Chemical Physics 1970;52:2523–2526.

[196] Turner R, von Kienlin M, Moonen C T, van Zijl P. Single-shot localized echo-planar imaging (steam-epi) at 4.7 tesla. Magnetic resonance in medicine 1990;14:401–408.

[197] Park H, Kim D, Cho Z. Gradient reversal technique and its applications to chemical-shift-related NMR imaging. Magnetic resonance in medicine 1987;4:526–536.

[198] Volk A, Tiffon B, Mispelter J, Lhoste J M. Chemical shift-specific slice selection. a new method for chemical shift imaging at high magnetic field. Journal of Magnetic Resonance (1969) 1987;71:168–174.

[199] Steidle G, Schick F. Echoplanar diffusion tensor imaging of the lower leg musculature using eddy current nulled stimulated echo preparation. Magnetic resonance in medicine 2006;55:541–548.

[200] Mattiello J, Basser P J, LeBihan D. Analytical expressions for the¡ i¿ b¡/i¿ matrix in NMR diffusion imaging and spectroscopy. Journal of magnetic resonance, Series A 1994;108:131–141.

[201] Walsh D, Gmitro A, Marcellin M. Adaptive reconstruction of phased array MR imagery. Magnetic Resonance in Medicine 2000;43:682–690.

[202] Dietrich O, Raya J, Reeder S, Ingrisch M, Reiser M, Schoenberg S. Influence of multichannel combination, parallel imaging and other reconstruction techniques on MRI noise characteristics. Magnetic resonance imaging 2008;26:754–762.

[203] Fieremans E, Jensen J H, Helpern J A. White matter characterization with diffusional kurtosis imaging. Neuroimage 2011;58:177–188.

[204] Horsfield M A, Barker G J, McDonald W I. Self-diffusion in CNS tissue by volume-selective proton NMR. Magnetic resonance in medicine 1994;31:637–644.

[205] Clark C A, Hedehus M, Moseley M E. Diffusion time dependence of the apparent diffusion tensor in healthy human brain and white matter disease. Magnetic resonance in medicine 2001;45:1126–1129.

[206] Ford J C, Hackney D B. Numerical model for calculation of apparent diffusion coefficients (adc) in permeable cylinders—comparison with measured adc in spinal cord white matter. Magnetic resonance in medicine 1997;37:387–394.

[207] Ford J C, Hackney D B, Lavi E, Phillips M, Patel U. Dependence of apparent diffusion coefficients on axonal spacing, membrane permeability, and diffusion time in spinal cord white matter. Journal of Magnetic Resonance Imaging 2005;8:775–782.

[208] Nair G, Tanahashi Y, Low H P, Billings Gagliardi S, Schwartz W J, Duong T Q. Myelination and long diffusion times alter diffusion-tensor-imaging contrast in myelin-deficient¡ i¿ shiverer¡/i¿ mice. Neuroimage 2005;28:165–174.

[209] Rane S, Nair G, Duong T Q. Dti at long diffusion time improves fiber tracking. NMR in Biomedicine 2010;23:459–465.

[210] Wang B, Kuo J, Bae S C, Granick S. When brownian diffusion is not gaussian. Nature Materials 2012;11:481–485.

[211] Berry A C. The accuracy of the gaussian approximation to the sum of independent variates. Transactions of the American Mathematical Society 1941;49:122–136.

[212] Esseen C G. A moment inequality with an application to the central limit theorem. Scandinavian Actuarial Journal 1956;1956:160–170.

[213] Haus J W, Kehr K W. Diffusion in regular and disordered lattices. Physics Reports 1987;150:263–406.

[214] Magin R L, Abdullah O, Baleanu D, Zhou X J. Anomalous diffusion expressed through fractional order differential operators in the bloch–torrey equation. Journal of Magnetic Resonance 2008;190:255–270.

[215] Jones D, Horsfield M, Simmons A. Optimal strategies for measuring diffusion in anisotropic systems by magnetic resonance imaging. Magn Reson Med 1999;42.

[216] Anderson A W. Theoretical analysis of the effects of noise on diffusion tensor imaging. Magnetic Resonance in Medicine 2001;46:1174–1188.

[217] Farrell J A, Landman B A, Jones C K, Smith S A, Prince J L, van Zijl P, Mori S. Effects of signal-to-noise ratio on the accuracy and reproducibility of diffusion tensor imaging–derived fractional anisotropy, mean diffusivity, and principal eigenvector measurements at 1.5 t. Journal of Magnetic Resonance Imaging 2007;26:756–767.

[218] Kakalios J, Street R, Jackson W. Stretched-exponential relaxation arising from dispersive diffusion of hydrogen in amorphous silicon. Physical review letters 1987;59:1037.

[219] Ozarslan E, Basser P J, Shepherd T M, Thelwall P E, Vemuri B C, Blackband S J. Observation of anomalous diffusion in excised tissue by characterizing the diffusion-time dependence of the MR signal. Journal of Magnetic Resonance 2006;183:315–323.

[220] Pasternak O, Sochen N, Gur Y, Intrator N, Assaf Y. Free water elimination and mapping from diffusion MRI. Magnetic Resonance in Medicine 2009;62:717–730.

[221] Karger J, Heink W. The propagator representation of molecular transport in microporous crystallites. Journal of Magnetic Resonance (1969) 1983;51:1–7.

[222] Wedeen V J, Hagmann P, Tseng W Y I, Reese T G, Weisskoff R M. Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging. Magnetic Resonance in Medicine 2005;54:1377–1386.

[223] Bennett K M, Schmainda K M, Rowe D B, Lu H, Hyde J S, et al. Characterization of continuously distributed cortical water diffusion rates with a stretched-exponential model. Magnetic resonance in medicine 2003;50:727–734.

[224] Henkelman R M. Measurement of signal intensities in the presence of noise in MR images. Medical physics 1985;12:232–233.

[225] Jones D, Basser P. "Squashing peanuts and smashing pumpkins": How noise distorts diffusion-weighted MR data. Magnetic Resonance in Medicine 2004;52:979–993.

[226] Sijbers J, Den Dekker A, Scheunders P, Van Dyck D. Maximum-likelihood estimation of rician distribution parameters. Medical Imaging, IEEE Transactions on 1998;17:357–361.

[227] Sijbers J, Den Dekker A. Maximum likelihood estimation of signal amplitude and noise variance from MR data. Magnetic Resonance in Medicine 2004;51:586–594.

[228] Zhang H, Schneider T, Wheeler Kingshott C A, Alexander D C. Noddi: Practical in vivo neurite orientation dispersion and density imaging of the human brain. Neuroimage 2012;61:1000–1016.

[229] Leergaard T B, White N S, de Crespigny A, Bolstad I, DArceuil H, Bjaalie J G, Dale A M. Quantitative histological validation of diffusion MRI fiber orientation distributions in the rat brain. PloS one 2010;5:e8595.

[230] Sehy J V, Ackerman J J, Neil J J. Evidence that both fast and slow water adc components arise from intracellular space. Magnetic resonance in medicine 2002;48:765–770.

[231] Schwarcz A, Bogner P, Meric P, Correze J L, Berente Z, Pal J, Gallyas F, Doczi T, Gillet B, Beloeil J C. The existence of biexponential signal decay in magnetic resonance diffusion-weighted imaging appears to be independent of compartmentalization. Magnetic resonance in medicine 2004;51:278–285.

[232] Nicholson C, Sykova E. Extracellular space structure revealed by diffusion analysis. Trends in neurosciences 1998;21:207–215.

[233] Kwong K, McKinstry R, Chien D, Crawley A, Pearlman J, Rosen B. Csf-suppressed quantitative single-shot diffusion imaging. Magnetic resonance in medicine 1991;21:157–163.

[234] Rooney W D, Johnson G, Li X, Cohen E R, Kim S G, Ugurbil K, Springer C S. Magnetic field and tissue dependencies of human brain longitudinal 1H2o relaxation in vivo. Magnetic Resonance in Medicine 2007;57:308–318.

[235] Landman B, Bazin P, Prince J. Diffusion tensor estimation by maximizing Rician likelihood. In: *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*;pages 1–8. IEEE 2007.

[236] Veraart J, Van Hecke W, Sijbers J. Constrained maximum likelihood estimation of the diffusion kurtosis tensor using a rician noise model. Magnetic Resonance in Medicine 2011.

[237] Kristoffersen A. Statistical assessment of non-gaussian diffusion models. Magnetic Resonance in Medicine 2011.

[238] Kristoffersen A. Estimating non-gaussian diffusion model parameters in the presence of physiological noise and rician signal bias. Journal of Magnetic Resonance Imaging 2012.

[239] Behrens T, Woolrich M, Jenkinson M, Johansen Berg H, Nunes R, Clare S, Matthews P, Brady J, Smith S. Characterization and propagation of uncertainty in diffusion-weighted MR imaging. Magnetic resonance in medicine 2003;50:1077–1088.

[240] Clarke R, Scifo P, Rizzo G, DellAcqua F, Scotti G, Fazio F. Noise correction on rician distributed data for fibre orientation estimators. Medical Imaging, IEEE Transactions on 2008;27:1242–1251.

[241] Kruger G, Glover G. Physiological noise in oxygenation-sensitive magnetic resonance imaging. Magnetic Resonance in Medicine 2001;46:631–637.

[242] Yacoub E, De Moortele V, Shmuel A, Ugurbil K, et al. Signal and noise characteristics of hahn se and ge bold fMRI at 7 t in humans. Neuroimage 2005;24:738–750.

[243] Hutton C, Josephs O, Stadler J, Featherstone E, Reid A, Speck O, Bernarding J, Weiskopf N. The impact of physiological noise correction on fMRI at 7t. Neuroimage 2011;57:101–112.

[244] Andersson L, Bolling M, Wirestam R, Holtas S, Stahlberg F. Combined diffusion weighting and csf suppression in functional MRI. NMR in Biomedicine 2002;15:235–240.

[245] Huang R, Chavez I, Taute K M, Lukic B, Jeney S, Raizen M G, Florin E L. Direct observation of the full transition from ballistic to diffusive brownian motion in a liquid. Nature Physics 2011;7:576–580.

[246] Chin C L, Wehrli F W, Hwang S N, Takahashi M, Hackney D B. Biexponential diffusion attenuation in the rat spinal cord: computer simulations based on anatomic images of axonal architecture. Magnetic resonance in medicine 2002;47:455–460.

[247] Fieremans E, Novikov D S, Jensen J H, Helpern J A. Monte carlo study of a two-compartment exchange model of diffusion. NMR in Biomedicine 2010;23:711–724.

[248] Nilsson M, Latt J, Nordh E, Wirestam R, Stahlberg F, Brockstedt S. On the effects of a varied diffusion time in vivo: is the diffusion in white matter restricted? Magnetic resonance imaging 2009;27:176–187.

[249] Badaut J, Lasbennes F, Magistretti P J, Regli L. Aquaporins in brain: distribution, physiology, and pathophysiology. Journal of Cerebral Blood Flow & Metabolism 2002;22:367–378.

[250] Amiry Moghaddam M, Ottersen O P. The molecular basis of water transport in the brain. Nature Reviews Neuroscience 2003;4:991–1001.

[251] Ellisman M H, Porter K R. Microtrabecular structure of the axoplasmic matrix: visualization of cross-linking structures and their distribution. The Journal of cell biology 1980;87:464–479.

[252] Avram L, Assaf Y, Cohen Y. The effect of rotational angle and experimental parameters on the diffraction patterns and micro-structural information obtained from q-space diffusion NMR: implication for diffusion in white matter fibers. Journal of Magnetic Resonance 2004;169:30–38.

[253] Huang H, Zhang J, Wakana S, Zhang W, Ren T, Richards L J, Yarowsky P, Donohue P, Graham E, Van Zijl P, et al. White and gray matter development in human fetal, newborn and pediatric brains. Neuroimage 2006;33:27–38.

[254] Nilsson M, Hagslatt H, Van Westen D, Wirestam R, Stahlberg F, Latt J. A mechanism for exchange between intraaxonal and extracellular water: Permeable nodes of ranvier. In: *Proc Intl Soc Mag Reson Med*;volume 18;page 1570 2010.

[255] Mikelberg F S, Drance S M, Schulzer M, Yidegiligne H M, Weis M M. The normal human optic nerve: axon count and axon diameter distribution. Ophthalmology 1989;96:1325–1328.

[256] Chomiak T, Hu B. What is the optimal value of the g-ratio for myelinated fibers in the rat CNS? a theoretical approach. PLoS One 2009;4:e7754.

[257] Alexander D C, Hubbard P L, Hall M G, Moore E A, Ptito M, Parker G J, Dyrby T B. Orientationally invariant indices of axon diameter and density from diffusion MRI. Neuroimage 2010;52:1374–1389.

# Abbreviations

| | |
|---|---|
| **ADC** | Apparent Diffusion Coefficient |
| **BwB** | Biexponential with Baseline |
| **CPMG** | Carr-Purcell Meiboom Gill |
| **CSF** | Cerebro-Spinal Fluid |
| **CHARMED** | Composite Hindered And Restricted Model of Diffusion |
| **CDF** | Cumulative Distribution Function |
| **DSI** | Diffusion Spectrum Imaging |
| **DTI** | Diffusion Tensor Imaging |
| **DWI** | Diffusion-Weighted Imaging |
| **DKI** | Diffusionl Kurtosis Imaging |
| **EPI** | Echo-Planar Imaging |
| **EAP** | Ensemble Average Propagator |
| **FLAIR** | Fluid Attenuation Inversion Recovery |
| **FA** | Fractional Anisotropy |
| **FID** | Free Induction Decay |
| **FWHM** | Full-Width at Half Maximum |
| **fMRI** | functional Magnetic Resonance Imaging |
| **GM** | Gray Matter |
| **HARDI** | High Angular Resolution Diffusion Imaging |
| **ILT** | Inverse Laplace Transform |
| **jMLE** | Joint-Maximum Likelihood Estimate |
| **MRI** | Magnetic Resonance Imaging |
| **MLE** | Maximum Likelihood Estimate |
| **MwB** | Monoexponential with Baseline |
| **NMR** | Nuclear Magnetic Resonance |
| **PDF** | Probability Density Function |

| | |
|---|---|
| **PFG** | Pulsed Field-Gradient |
| **PGSE** | Pulsed Gradient Spin Echo |
| **SNR** | Signal-to-Noise Ratio |
| **SAR** | Specific Absorption Rate |
| **SE** | Spin Echo |
| **STEAM** | Stimulated Echo Acquisition Mode |
| **TrSe** | Twice-refocused Spin Echo |
| **WM** | White Matter |
| **MBP** | Myelin Basic Protein |
| **PLP** | ProtoLipid protein |
| **MAG** | Myelin Association Glycoprotein |

# List of Figures

# List of Tables

# *Acknowledgements*

First of all I would like to thank Professor Turner for his acceptance, support and guidance throughout these years. For me who started from the basics of MRI and zero knowhow of biological system, it was a wonderful chance to learn directly from somebody who knew both the subject matter in depth. Throughout our discussions, I always felt that I came back with a lot of new information and a different understanding of how things work. His enthusiasm in a variety of topics have always fascinated me and kept me pushing to learn more. The freedom during discussions and freedom of ideas allowed me to choose my research direction but also made me feel more responsible towards the output of my work.

Professor Möller has been very warm to me throughout these years - always willing to help if I needed something. He was there at a distance but always approachable. Dr. Christian Labadie has helped me a lot in a range of things - writing, programming and thinking about brain structures.

I would like to thank Professor Kärger, Dr. Galvosas and Dr. Stallmach who allowed me to measure in the MR spectrometers especially designed for diffusion measurements. Many measurements that were instrumental in my understanding of diffusion as well as NMR were done there. Marcel Gratz was very helpful in teaching me how to measure in the spectrometer and to show me what were the new advance techniques in the field. I am thankful to Professor Käs and Undine for letting me use their cell culture facility during the early days. I had started out with an ambitious desire to grow neurons to be able to measure diffusion in growing neuronal processes. Little did 'i' know that cultured neurons are as sensitive as neurons *in vivo* and keeping their processes intact for measurements turned out to be a huge unsuccessful task. The experience taught me a valuable lesson - about the significance of these processes for neuronal development - without them the neurons were just like ordinary cells.

Many friends at the Max Planck Institute are crucial to my work and life. Robert Trampel was always helpful, I could knock on his door any time with a question or without a question. Andreas Schäfer's calmness almost belies his knowledge. In many situations when I have been bothered by even the tiniest of procedures during MR imaging - which happened very often during initial stages of sequence programming - I constantly needed his support. Markus, Dimo and Eugenia were nice and very supportive - never discouraging me for asking wide range of questions, and still supporting me with

corrections. Many thanks to Aline Peter who has been very helpful, without her it would have taken me months more to settle in Leipzig and months more to get out of there! Talking to Aline is always a pleasure. I am also grateful to Enrico Reimer who helped me throughout with computer related problems and data conversion into different format. Alfred Anwander introduced me to diffusion modeling and analysis. And thank you Pierre-Louis for showing me how images can be morphed, registered and segmented.

I would also like to thank my friend Subhasis for his constant support: giving me a place to crash, reminding me of deadlines and constantly listening to my 'new' ideas that were remotely related to his field of work.

I would also like to thank Bidhi and Anam for helping me with the corrections.

Finally, I would like to thank my dad and my mom who were very encouraging and always happy to hear from me. Although they live far away and did not get to see my transformation during the course of this work, love has always been....

...